前言

最近几年来，地理信息系统无论是在理论上还是应用上都处在一个飞速发展的阶段。GIS 被应用于多个领域的建模和决策支持，如城市管理、区划、环境整治等等，地理信息成为信息时代重要的组成部分之一；“数字地球”概念的提出，更进一步推动了作为其技术支撑的 GIS 的发展。与此同时，一些学者致力于相关的理论研究，如空间感知、空间数据误差、空间关系的形式化等等。这恰好说明了地理信息系统作为应用技术和学科的两个方面，并且这两个方面构成了相互促进的发展过程。

从应用的角度看，近几年 GIS 的发展主要体现在以下几个方面：

1：应用领域的扩大，如一些商业部门开始利用 GIS 进行客户研究；此外同遥感、GPS 的结合，为 GIS 开拓了更加广泛的应用空间，如汽车导航，环境监测等。

2：大量成熟的商业 GIS 平台出现，或已有的软件不断推出新的产品。

3：空间数据的建设越来越受到重视，基于空间数据基础设施的建设，开始了空间数据共享和互操作的研究。

在这样的背景下，全国许多高校的地理系都开始了 GIS 专业人才的培养，以满足 GIS 发展的需求。北京大学城市与环境学系于 1999 年开始招收 GIS 专业本科生，这样，编写一本全面的、能够反映 GIS 最新进展的教材就成为当务之急。所以，作者编写这本书。

在编写的过程中，考虑到 GIS 专业学生今后从事的工作可能包括以下几类：

1：GIS 理论的研究工作；

2：GIS 软件的开发；

3：GIS 项目组织管理；

4：应用 GIS 的其它领域。

因此，在设计本书的结构时兼顾了这些方面，依次介绍了 GIS 的基本理论、功能实现、应用技术。本书共二十一章，分为五部分：

1：第一章，介绍了 GIS 的相关概念，信息、数据、信息系统，以及 GIS 的发展历史和展望；

2：第二——五章，介绍了 GIS 的空间数据模型、参照系统和 GIS 中的数据，这几章组成了 GIS 的理论基础；

3：第六——十一章，按照从空间数据录入、管理、分析、表现的过程讲述了 GIS 的基本功能；

4：第十二——十九章，介绍了 GIS 应用技术的各个方面。

5：附录（两章），分别介绍了 GIS 的计算机基础和国内外的 GIS 工具软件，供阅读本书时参考。

在编写中，每章前面的导读概括了本章的内容，说明了与其它章节的关系；正文中采用灰底文字讲述了相关的知识，供阅读时参考。

本书作为教材而编写，但同时希望它能够对 GIS 开发人员、管理人员有所帮助。

本书由邬伦、刘瑜确定整体结构，主要编写人员有刘瑜、张晶、马修军、韦中亚、田原等。各章主笔分工为：第一——五章，第七章，第十一章：张晶；第八、九、十章：马修军；第六章和附录一：韦中亚；第十二章——第十九章，附录二：刘瑜。最终本书由刘瑜、邬伦统稿，邬伦定稿。

本书的完成，是北京大学城市与环境学系 GIS 教研室的多年工作的结晶，教研室人员的工作，包括 Citystar 软件的开发以及相关的科研工作，都为本书提供了极好的“原料”。教研室中，下列人员参加了书稿的讨论以及部分文字整理工作：叶燕林，张金海，刘岳峰，

林报嘉，唐大仕，高勇，张毅，王永乾，周贵云，苏元峰，刘玉玲，赵建伟，刘述，赵伟。

本书的完成，得到了许多老师和同行的关怀和支持。德高望重的陈述彭院士于百忙之中为本书作序；何建邦教授、承继成教授、谢昆青博士、任伏虎博士、张世琨博士为本书提出了许多有益的建议和修改意见；山西大学马义娟，江西省气象局曹华盖、文海龙也参加了部分工作，在此表示衷心的感谢。

地理信息系统一门实用性很强的学科，必须更多地动手实践，才能更好地理解其中的一些概念，因此与本书配套，作者还将继续编写地理信息系统实习教材，并开发相应的实习软件。

由于作者水平和时间所限，书中错误在所难免，希望读者不吝指正。

作者
2000/9/1

序

国务院学位委员会决定：新生的地理信息系统与古老的地图学并列为理科专业。在科学出版社新版的《北京大学地理教学丛书》中，邬伦教授等主编的《地理信息系统——原理、方法和应用》又列为其中系列教材之一，同时也作为大学本科、研究生及自学考研的参考书。它的出版，标志着地理信息系统作为大学理科专业已经到了科技界、教育界的认同——它既是应用领域广泛的工程技术；又是跨越信息科学、地球科学和空间科学的应用基础科学。这是科教兴国战略与教育深化改革的又一崭新的亮点。我感到非常欣慰，并表示热烈的祝贺！

随着全球化、网络化和知识经济蓬勃发展的新世纪的到来，我国地理信息系统正迅猛地由实验室走向产业化和社会化。由设备引进、技术克隆迈上自主开发、数据挖掘与知识创新的新台阶。随着加入世界贸易组织（WTO），在互联网服务领域（ISP）、信息服务领域（ICP）、以在用服务领域（ASP），都需要大量地理信息系统（GIS）人才，特别是管理科学结合的管理地理信息系统（MGIS）人才。

我国地理信息系统起步稍晚。20 世纪末，互联网的普及还远远低于世界平均水平。计算机软、硬件人员在信息产业中所占比例低于 15%。虽然已有 74%的重点企业建立了网站主页，但大部分还没有建立防火墙设施，安全审计、监测和自动恢复系统。在国家重点企业中，采用的服务器惠普、IBM 各占 35%以上，UPS 山特独占 69%，路由器 CISCO 占 70%以上，交换机 3COM 占 42%，操作系统 WINDOW 占 51%，办公软件 OFFICE 占 83.3%（据国家经贸委经济信息中心调查）。在迎接全球化、网络化新机遇的同时，我国信息化所面临的挑战也是十分严峻的，任重而道远，关键在于素质教育与科技人才的培养。我国联想生产的 PC 机占有国内市场的 55%，北京大学的方正系统、Citystar 系统和其他院校开发的软件已享有盛誉，其中地理信息系统中自主开发的软件已占有国内市场 25%左右，是非常良好的开端，但在发达国家中，信息产业对国民经济增长贡献率和就业动力所占社会劳动力都已超过 50%，我们的成就还只能算是万里长征的第一步。

近年出版了一系列地理信息系统教材，其中有些侧重于科学原理阐明，有些侧重于专业部门的应用。邬伦博士以其深厚的地学造诣与计算机软件开发的丰富实践，融汇贯通化为一体，又着眼于公共课程的广大性与大学本科的基础性，旁征博引，求实创新，图文并茂，深入浅出，推出这部迎接新世纪空间时代和信息社会的好教材，是非常难能可贵的。我是非常乐意向读者们郑重推荐的。

陈述彭

2000 年 10 月 12 日

《地理信息系统原理、方法和应用》书稿评审意见

关于地理信息系统的书籍、至少已经出版了数十本之多，但是由北京大学邬伦教授最近编写的《地理信息系统原理、方法和应用》书稿具有与众不同的特点，读完之后，使人有新鲜的感觉，该书的主要优点是：不仅正确而又精练地介绍了地理信息系统的基本内容，而且又把本学科的最新概念，最新成果和动态趋势作了全面、系统的介绍，如从现实世界到比特世界（第二章），35 集成技术（第十二章），网络地理信息系统（第十三章），地理信息系统应用项目组织和管理（第十五章），地理信息系统软件工程技术（第十六章），地理信息系统标准（第十七章），地理信息系统和社会（第十八章），地球信息科学与数字地球（第十七章），地理信息系统和社会（第十八章），地球信息科学与数字地球（第十九章）和 GIS 的计算基础（第二十章）等，不仅是必要的，而且也是合适的。

本书的另一个特点是它的结构合理，逻辑性强，文笔通顺，图文并茂，深入浅出，可读性好。

该书不仅适合作大专院校的教材以及补充读物，也适合给工程技术人员，科研工作者和其他专业人员作为参考书用，是一本很好的介绍地理信息系统的著作。

建议及早将它出版，与广大读者见面，估计定能受到欢迎好评。

北京大学遥感与 GIS 研究所教授、博导
国际欧亚科学院院士
承继成
2000 年 9 月 24 日

概论

1. GIS的定义，GIS是什么或者不是什么
2. 数据、信息、信息系统
3. GIS的发展历史
4. GIS功能概述(A Overview)
5. GIS和其他学科的关系
   - 概述
   - AM/FM
   - LIS
6. GIS发展展望

空间信息和空间数据

1. 现实世界到比特世界
2. 空间数据模型
   - 要素模型
   - 场模型
   - 网络结构
   - 时间
   - 三维
3. 空间参照系
   - 空间参照系
   - 地图投影
4. GIS中的数据
   - 数据类型
   - 数据涵义
   - 数据质量
   - 元数据

GIS的功能

1. 计算机基础
    - 硬件
    - 数据库
    - 算法和数据结构
    - 面向对象技术
2. 空间数据获取
    - 数字化
    - 数据规范处理
3. 空间数据管理
    - 栅格数据
    - 矢量数据
    - 矢量-栅格转换
    - 空间索引技术
    - 空间查询和检索
4. 空间分析
    - 空间关系检测
    - 空间度量
    - 空间切分
    - 网络分析
    - Buffer
    - Overlay
5. 空间分析（二）
    - 实例
    - 资源配置模型
    - 知识挖掘
    - 决策支持模型
    - 时间序列模型
6. DEM
    - DEM和DTM

- 插值
- DEM应用

7. 空间数据表现
   - 符号
   - 专题信息表现
   - 地图设计
   - 空间信息可视化
   - 制图综合

GIS应用(Applications)

1. GIS应用举例
   - 城市规划
   - 城镇管理
   - 电力
   - 交通
   - 环境
   - 农业
   - 林业
   - 地质地貌
   - 气象
2. GIS平台介绍
   - ARC/INFO
   - MapInfo
   - Intergraph
   - 国产GIS
3. 3S集成
   - RS介绍
   - GPS介绍
   - 3S集成应用

4. GIS项目以及项目组织
    - GIS项目类型（按照目的，规模进行分类）
    - 资金
    - 人员
    - 数据
    - 进度控制
5. GIS软件工程
    - 基础软件开发
    - 应用系统开发
    - 质量控制质量标准
6. 网络GIS
    - 计算机网络技术
    - 分布式空间数据管理
    - WebGIS
7. GIS的标准化
8. GIS与社会
    - 走出学院的GIS
    - 社会对GIS发展的影响
    - GIS与社会，另外几个问题

      教育

      产业

      版权
9. 数字地球

如果说地图是地理学的第二代语言，那么地理信息系统就是地理学的第三代语言。

陈述彭

简单的说，地理信息系统是回答如下问题的信息系统——“哪里”。

Henry Tom

第一章 地理信息系统概论

导读：本章介绍了地理信息系统的一些最基本的、但又是非常重要的概念，包括信息、数据、信息系统、空间数据、空间信息和地理信息系统。有些概念，如空间信息和空间数据，还将得到进一步的阐述，实际上，空间数据也是本书中描述 GIS 功能的核心。

GIS 既是一项技术，也是一门学科，本章也介绍了它的主要研究内容以及与其它学科，如地理学、地图学、遥感等的关系。

本章同时介绍了 GIS 的组成、分类、功能、发展历史和展望，这些将在后面的章节中更为详尽地被描述。

1．地理信息系统的基本概念

1． 1 信息、数据、地理数据与地理信息

1．1．1 信息（Information）

1）信息的含义

信息是近代科学的一个专门术语，已广泛地应用于社会各个领域。狭义信息论将信息定义为“两次不定性之差”，即指人们获得信息前后对事物认识的差别；广义信息论认为，信息是指主体（人、生物或机器）与外部客体（环境、其他人、生物或机器）之间相互联系的一种形式，是主体和客体之间的一切有用的消息或知识，是表征事物特征的一种普遍形式。本书采用的定义为：信息是向人们或机器提供关于现实世界新的事实的知识，是数据、消息中所包含的意义，它不随载体物理设备形式的改变而改变。

2）信息的特点

信息具有以下特点：

- 客观性：任何信息都是与客观事实紧密相关的，这是信息正确性和精确度的保证；
- 实用性：信息对决策是十分重要的，信息系统将地理空间的巨大数据流收集、组织和管理起来，经过处理、转换和分析变为对生产、管理和决策具有重要意义的有用信息；
- 传输性：信息可以在信息发送者和接受者之间传输，既包括系统把有用信息送至终端设备（包括远程终端）和以一定的形式或格式提供给有关用户，也包括信息在系统内各个子系统之间的流转和交换，如网络传输技术；
- 共享性：信息与实物不同，信息可以传输给多个用户，为多个用户共享，而其本身并无损失。信息的这些特点，使信息成为当代社会发展的一项重要资源。

1．1．2 数据

数据是指某一目标定性、定量描述的原始资料，包括数字、文字、符号、图形、图像以及它们能转换

成的数据等形式。数据是用以载荷信息的物理符号，数据本身并没有意义。如数字“1”。信息可以离开信息系统而独立存在，也可以离开信息系统的各个组成和阶段而独立存在；而数据的格式往往与计算机系统有关，并随载荷它的物理设备的形式而改变。

信息与数据是不可分离的。信息由与物理介质有关的数据表达，数据中所包含的意义就是信息。数据是记录下来的某种可以识别的符号，具有多种多样的形式，也可以由一种数据形式转换为其他数据形式，但其中包含的信息的内容不会改变。数据是信息的载体，但并不就是信息。只有理解了数据的含义，对数据做出解释，才能提取数据中所包含的信息。对数据进行处理（运算、排序、编码、分类、增强等）就是为了得到数据中包含的信息。虽然日常生活中数据和信息概念分得不是很清，但它们有着不同的含义。可以把数据比作原材料，而信息是对原材料处理的结果。如同一个木匠，在一些工具的帮助下，可以把木材做成有用的家具。同样，计算机专业人员应用计算机的硬件和软件把原始数据转换成信息。这种转换过程可用图 1-1 说明。

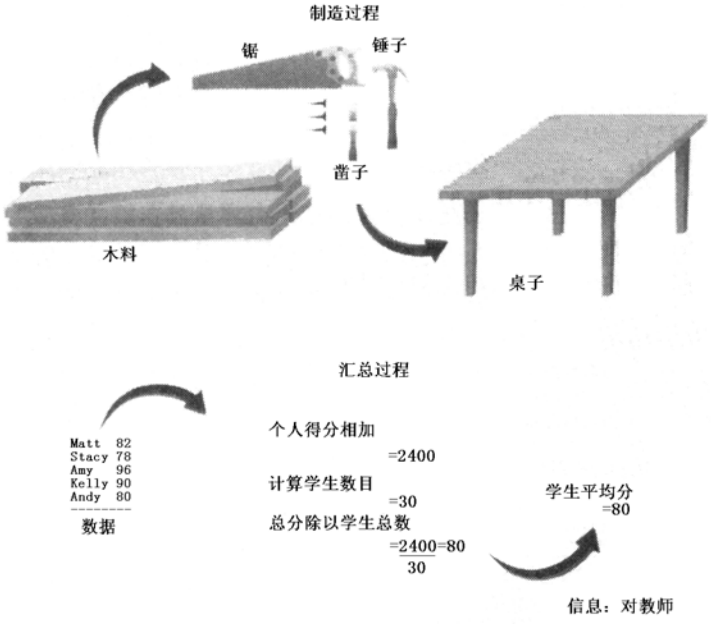


图 1-1：数据和信息

数据包含原始事实，信息是把数据处理成有意义的和有用的形式。例如，每个学生所得分数对教师来说是原始数据，而把这些数据进行汇总并计算每班平均得分，教师根据平均分判断班级总体情况，这些结果对教师来说就是信息。与数据相比，信息具有以下特征：数据是原始事实，信息是数据处理的结果；对一个人是信息对其他人可能是数据；信息必须是有意义或有用的；使用的信息必须是完整、精确、相关和及时的。

人的知识、经验作用到数据上，可以得到信息，而获得信息量的多少，与人的知识水平有关[*]。

1．1．3 地理信息和地理数据

地理数据 是指表征地理圈或地理环境固有要素或物质的数量、质量、分布特征、联系和规律的数字、文字、图像和图形等的总称。地理信息是有关地理实体的性质、特征和运动状态的表征和一切有用的知识，它是对地理数据的解释。在地理信息中，其位置是通过数据进行标识的，这是地理信息区别于其它类型信息的最显著的标志。地理信息具有区域性、多维结构特性和动态变化的特性：

1）区域性是通过经纬网等建立的地理坐标来实现空间位置的标识；

2）多维结构特性即在二维空间的基础上实现多专题的第三维结构；

---

[*] 以上述的制作桌子的例子说明，同样的木材（数据），由于使用工具不同和制作者水平的高低，制作的桌子（信息）也不同。

3）地理信息的时序特征十分明显，可以按时间尺度将地理信息划分为超短期的（如台风、地震）、短期的（如江河洪水、秋季低温）、中期的（如土地利用、作物估产）、长期的（如城市化、水土流失）、超长期的（如地壳变动、气候变化）等。

地理数据是各种地理特征和现象间关系的符号化表示，包括空间位置、属性特征及时态特征三部分。空间位置数据描述地物所在位置，这种位置既可以根据大地参照系定义，如大地经纬度坐标，也可以定义为地物间的相对位置关系，如空间上的距离、邻接、重叠、包含等；属性数据又称为非空间数据，是属于一定地物、描述其特征的定性或定量指标，即描述了信息的非空间组成部分，包括语义与统计数据等；时态特征是指地理数据采集或地理现象发生的时刻或时段，时态数据对环境模拟分析非常重要，越来越受到地理信息系统学界的重视。从地理实体到地理数据、从地理数据到地理信息的发展，反映了人类认识的一个巨大飞跃。

1．2 信息系统（Information System）及其类型

系统（System）是具有特定功能的、相互有机联系的许多要素所构成的一个整体。对计算机而言，系统是为实现某些特定的功能，由必要的人、机器、方法或程序按一定的相关关系联系起来进行工作的集合体，内部要素之间的相互联系通过信息流实现。系统的特征由构成系统的要素及其相互之间的联系方式所决定。

1．2．1 信息系统

信息系统是具有数据采集、管理、分析和表达数据能力的系统，它能够为单一的或有组织的决策过程提供有用的信息。在计算机时代，信息系统都部分或全部由计算机系统支持，人们常常使用计算机收集数据并将数据处理成信息，计算机的使用导致了一场信息革命，目前，计算机已经渗透到各个领域。一个基于计算机的信息系统包括计算机硬件、软件、数据和用户四大要素。

1）计算机硬件包括各类计算机处理及终端设备，它帮助人们在非常短的时间内处理大量数据、存储信息和快速获得帮助；

2）软件是支持数据信息的采集、存储加工、再现和回答用户问题的计算机程序系统，它接受有效数据，并正确地处理数据；在一定的时间内提供适用的、正确的信息；并存储信息为将来所用；

3）数据是系统分析与处理的对象，构成系统的应用基础；

4）用户是信息系统所服务的对象。由于信息系统并不是完全自动化的，在系统中总是包含一些人的复杂因素，人的作用是输入数据、使用信息和操作信息系统，建立信息系统也需要人的参与。

在基于计算机的信息系统中，处理过程的作用是告诉人们各部分间的相互关系（图 1-2）。

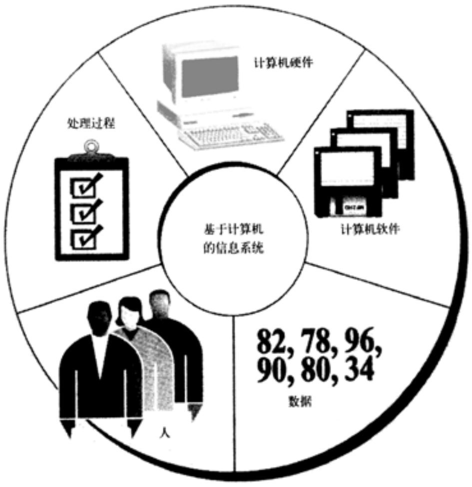


图 1-2：信息系统的组成

1．2．2 信息系统的类型

信息的需要完全取决于管理的层次，设计一个系统要满足组织中所有层次人员的信息需要，这种系统是很复杂的，因为组织中使用的信息在数量、状态和类型上都是易变和不可预知的。在组织中将信息系统分成三个管理层次：操作层（底层）、战术层（中间层）和战略层（顶层）。操作层包括的人员如会计师、销售人员和商店监理，他们执行日常工作和上级管理所做的计划；战术层包括组织中的高级管理人员和参与最高管理的中层管理人员；而管理层负责决定组织的发展方向。为了解决系统复杂性这一问题，大多数组织建立不同类型的系统来满足他们的需要，见信息系统类型图（图 1-3）：

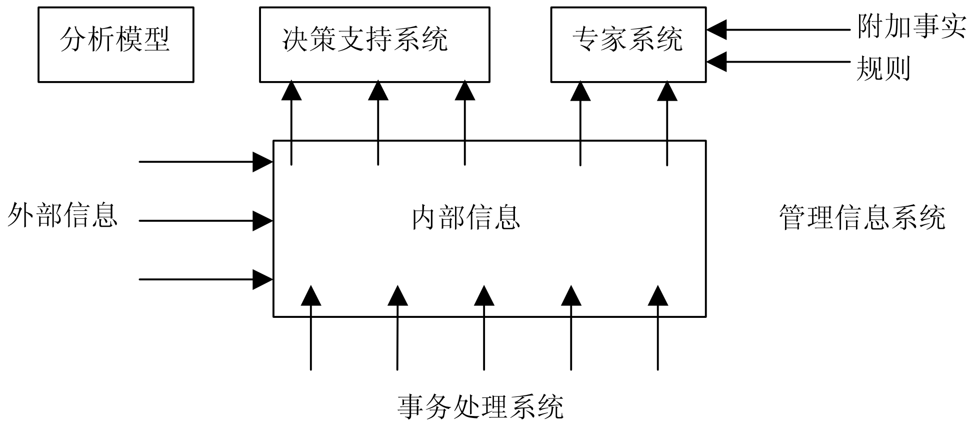


图 1-3：信息系统的类型

1）事物处理系统（Transaction Process System ，TPS）主要用以支持操作层人员的日常活动。它主要负责处理日常事务。

2）管理信息系统（Management Information System ，MIS）需要包含组织中的事务处理系统，并提供了内部综合形式的数据，以及外部组织的一般范围和大范围的数据。许多战术层提供的信息能按照该层管理者希望的那样以熟悉的和喜欢的形式提供。但是，为战术层管理者提供的另外一部分信息和大多数为战略层管理者提供的信息是不可能事先确定的。这些不确定性对管理信息系统的设计者来说是个很大的挑战。

3）决策支持系统（Decision Support System ，DSS）能从管理信息系统中获得信息，帮助管理者制定好的决策。该系统是一组处理数据和进行推测的分析程序，用以支持管理者制定决策。它是基于计算机的交互式的信息系统，由分析决策模型、管理信息系统中的信息、决策者的推测三者相组合达到好的决策效果。

4）人工智能和专家系统 专家系统（Expert System ,ES）是能模仿人工决策处理过程的基于计算机的信息系统。专家系统扩大了计算机的应用范围，使其从传统的资料处理领域发展到智能推理上来。MIS 能提供信息帮助制定决策，DSS 能帮助改善决策的质量，只有专家系统能应用智能推理制作决策并解释决策理由。专家系统由五个部分组成：知识库、推理机、解释系统、用户接口和知识获得系统。

2．地理信息系统及其类型

2．1 地理信息系统

地理信息系统（Geographical Information System，GIS）是一种决策支持系统，它具有信息系统的各种特点。地理信息系统与其他信息系统的主要区别在于其存储和处理的信息是经过地理编码的，地理位置及

与该位置有关的地物属性信息成为信息检索的重要部分。在地理信息系统中，现实世界被表达成一系列的地理要素和地理现象，这些地理特征至少由空间位置参考信息和非位置信息两个组成部分。

地理信息系统的定义是由两个部分组成的。一方面，地理信息系统是一门学科*，是描述、存储、分析和输出空间信息的理论和方法的一门新兴的交叉学科；另一方面，地理信息系统是一个技术系统，是以地理空间数据库（Geospatial Database）为基础，采用地理模型分析方法，适时提供多种空间的和动态的地理信息，为地理研究和地理决策服务的计算机技术系统。

地理信息系统具有以下三个方面的特征：

第一，具有采集、管理、分析和输出多种地理信息的能力，具有空间性和动态性；

第二，由计算机系统支持进行空间地理数据管理，并由计算机程序模拟常规的或专门的地理分析方法，作用于空间数据，产生有用信息，完成人类难以完成的任务；

第三，计算机系统的支持是地理信息系统的重要特征，因而使得地理信息系统能以快速、精确、综合地对复杂的地理系统进行空间定位和过程动态分析。

地理信息系统的外观，表现为计算机软硬件系统；其内涵却是由计算机程序和地理数据组织而成的地理空间信息模型。当具有一定地学知识的用户使用地理信息系统时，他所面对的数据不再是毫无意义的，而是把客观世界抽象为模型化的空间数据，用户可以按应用的目的观测这个现实世界模型的各个方面的内容，取得自然过程的分析和预测的信息，用于管理和决策，这就是地理信息系统的意义。一个逻辑缩小的、高度信息化的地理系统，从视觉、计量和逻辑上对地理系统在功能方面进行模拟，信息的流动以及信息流动的结果，完全由计算机程序的运行和数据的变换来仿真。地理学家可以在地理信息系统支持下提取地理系统各不同侧面、不同层次的空间和时间特征，也可以快速地模拟自然过程的演变或思维过程的结果，取得地理预测或“实验”的结果，选择优化方案，用于管理与决策。

地理信息系统和 4M

对于地理学家、城市规划人员、资源管理者等使用地理信息的人们而言，他们观察并量测（Measure）环境参数，并制作地图（Map）描述了地球的特征；他们监测（Monitor）周围环境的变化，对影响环境的过程和活动建立模型（Model）。这四项活动（4M），通过应用 GIS 技术，都可以得到改善和提高。

2．2 地理信息系统的类型

地理信息系统按其内容可以分为三大类：

1）专题地理信息系统（Thematic GIS），是具有有限目标和专业特点的地理信息系统，为特定的专门目的服务。例如，森林动态监测信息系统、水资源管理信息系统、矿业资源信息系统、农作物估产信息系统、草场资源管理信息系统、水土流失信息系统等。

2）区域信息系统（Regional GIS），主要以区域综合研究和全面的信息服务为目标，可以有不同的规模，如国家级的、地区或省级的、市级和县级等为各不同级别行政区服务的区域信息系统；也可以按自然分区或流域为单位的区域信息系统。区域信息系统如加拿大国家信息系统、中国黄河流域信息系统等。许多实际的地理信息系统是介于上述二者之间的区域性专题信息系统，如北京市水土流失信息系统、海南岛土地评价信息系统、河南省冬小麦估产信息系统等。

3）地理信息系统工具或地理信息系统外壳（GIS Tools），是一组具有图形图像数字化、存储管理、查询检索、分析运算和多种输出等地理信息系统基本功能的软件包。它们或者是专门设计研制的，或者在完成了实用地理信息系统后抽取掉具体区域或专题的地理系空间数据后得到的，具有对计算机硬件适应性强、数据管理和操作效率高、功能强且具有普遍性的实用性信息系统，也可以用作 GIS 教学软件。

在通用的地理信息系统工具支持下建立区域或专题地理信息系统，不仅可以节省软件开发的人力、物力、财力，缩短系统建立周期，提高系统技术水平，而且使地理信息系统技术易于推广，并使广大地学工作者可以将更多的精力投入高层次的应用模型开发上。

---

* 在“地球信息系统和数字地球”一章中提到的“地理信息科学”更加强调学科特性，与之对应，GIS则作为应用技术。

2．3 地理信息系统的构成

与普通的信息系统类似，一个完整的 GIS 主要由四个部分构成，即计算机硬件系统、计算机软件系统、地理数据（或空间数据）和系统管理操作人员。其核心部分是计算机系统（软件和硬件），空间数据反映 GIS 的地理内容，而管理人员和用户则决定系统的工作方式和信息表示方式。系统构成如图 1-4 所示。

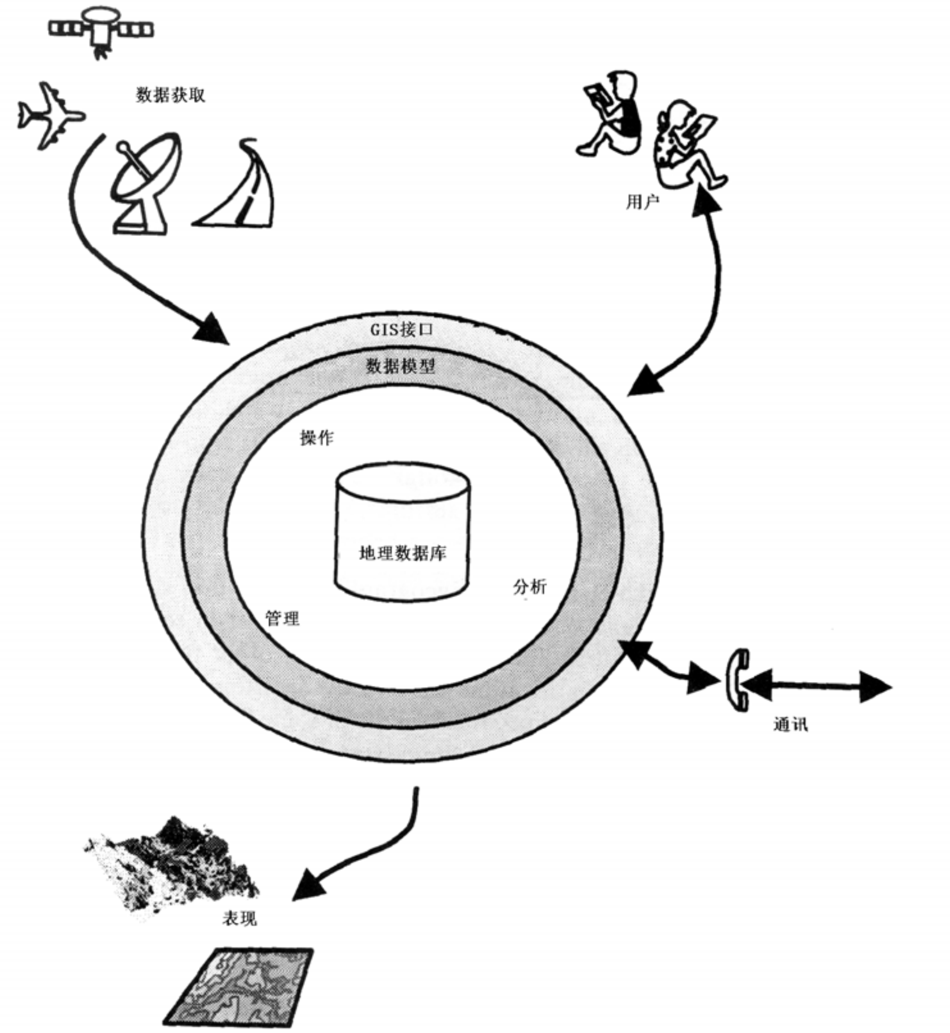


图 1-4：地理信息系统的构成

2．3．1 计算机硬件系统

计算机硬件系统是计算机系统中的实际物理装置的总称，可以是电子的、电的、磁的、机械的、光的元件或装置，是GIS的物理外壳。系统的规模、精度、速度、功能、形式、使用方法甚至软件都与硬件有极大的关系，受硬件指标的支持或制约。GIS由于其任务的复杂性和特殊性，必须由计算机设备支持。构成计算机硬件系统的基本组件包括输入/输出设备、中央处理单元、存储器（包括主存储器、辅助存储器硬件*）等，这些硬件组件协同工作，向计算机系统提供必要的信息，使其完成任务；保存数据以备现在或将来使用；将处理得到的结果或信息提供给用户。图 1-5 表示了常见的实现输入输出功能的计算机外部设备。

---

* 通常又称为内存和外存。

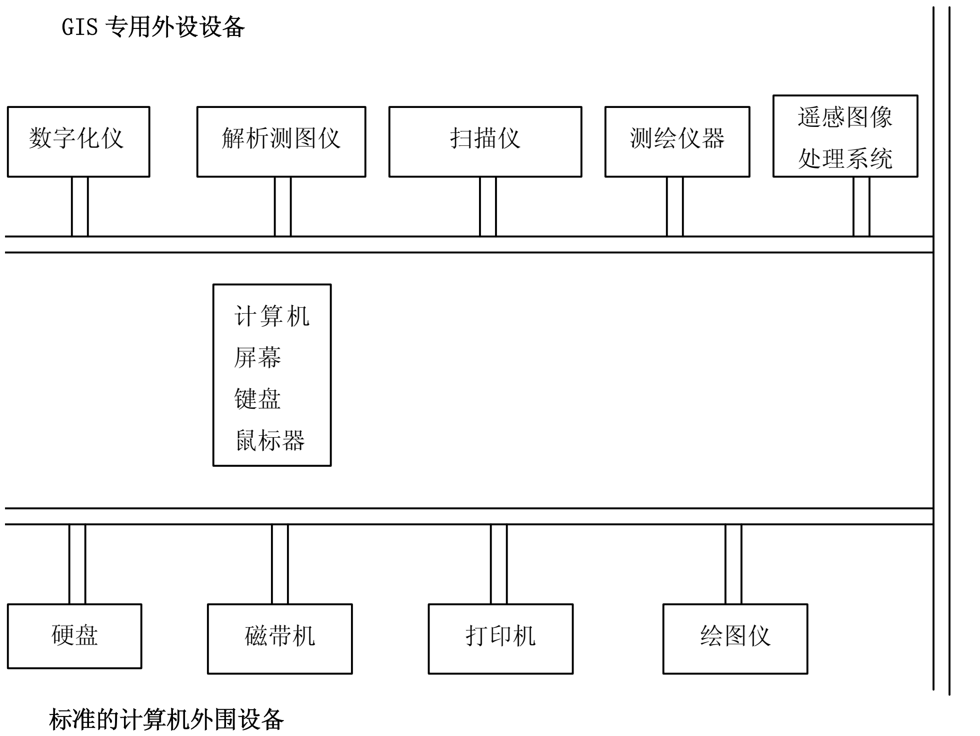


图 1-5：计算机标准外设和 GIS 使用的外设

3．3．2 计算机软件系统

计算机软件系统是指必需的各种程序。对于 GIS 应用而言，通常包括：

1）计算机系统软件

由计算机厂家提供的、为用户使用计算机提供方便的程序系统，通常包括操作系统、汇编程序、编译程序、诊断程序、库程序以及各种维护使用手册、程序说明等，是 GIS 日常工作所必需的。

2）地理信息系统软件和其他支持软件

包括通用的 GIS 软件包，也可以包括数据库管理系统、计算机图形软件包、计算机图像处理系统、CAD 等，用于支持对空间数据输入、存储、转换、输出和与用户接口。GIS 软件包功能结构见图 1-6。

3）应用分析程序

是系统开发人员或用户根据地理专题或区域分析模型编制的用于某种特定应用任务的程序，是系统功能的扩充与延伸。在 GIS 工具支持下，应用程序的开发应是透明的和动态的，与系统的物理存储结构无关，而随着系统应用水平的提高不断优化和扩充。应用程序作用于地理专题或区域数据，构成 GIS 的具体内容，这是用户最为关心的真正用于地理分析的部分，也是从空间数据中提取地理信息的关键。用户进行系统开发的大部分工作是开发应用程序，而应用程序的水平在很大程度上决定系统的应用性优劣和成败。

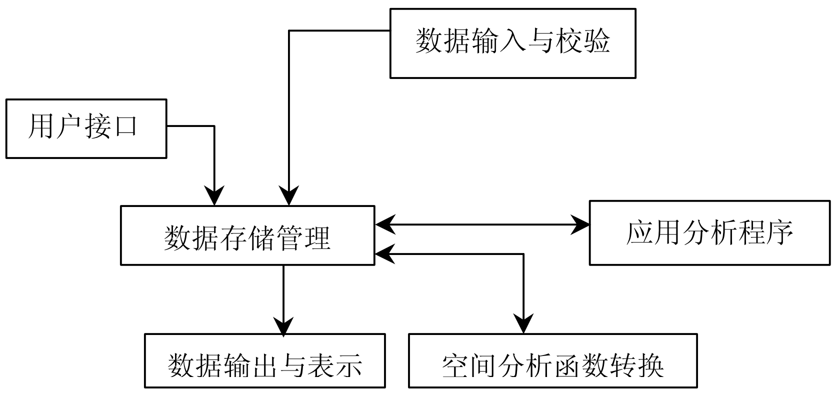


图 1-6：地理信息系统软件的功能框架

2．3．3 系统开发、管理和使用人员

人是 GIS 中的重要构成因素，GIS 不同于一幅地图，而是一个动态的地理模型。仅有系统软硬件和数据还不能构成完整的地理信息系统，需要人进行系统组织、管理、维护和数据更新、系统扩充完善、应用程序开发，并灵活采用地理分析模型提取多种信息，为研究和决策服务。对于合格的系统设计、运行和使用来说，地理信息系统专业人员是地理信息系统应用的关键，而强有力的组织是系统运行的保障。一个周密规划的地理信息系统项目应包括负责系统设计和执行的项目经理、信息管理的技术人员、系统用户化的应用工程师以及最终运行系统的用户。

2．3．4 空间数据*

是指以地球表面空间位置为参照的自然、社会和人文经济景观数据，可以是图形、图像、文字、表格和数字等。它是由系统的建立者通过数字化仪、扫描仪、键盘、磁带机或其他系统通讯输入 GIS，是系统程序作用的对象，是 GIS 所表达的现实世界经过模型抽象的实质性内容。在 GIS 中，空间数据主要包括：

1）某个已知坐标系中的位置

即几何坐标，标识地理景观在自然界或包含某个区域的地图中的空间位置，如经纬度、平面直角坐标、极坐标等，采用数字化仪输入时通常采用数字化仪直角坐标或屏幕直角坐标。

2）实体间的空间关系

实体间的空间关系通常包括：度量关系，如两个地物之间的距离远近；延伸关系（或方位关系），定义了两个地物之间的方位；拓扑关系，定义了地物之间连通、邻接等关系，是 GIS 分析中最基本的关系，其中包括了网络结点与网络线之间的枢纽关系（图 1-7-a），边界线与面实体间的构成关系（图 1-7-b），面实体与岛或内部点的包含关系（图 1-7-c）等。

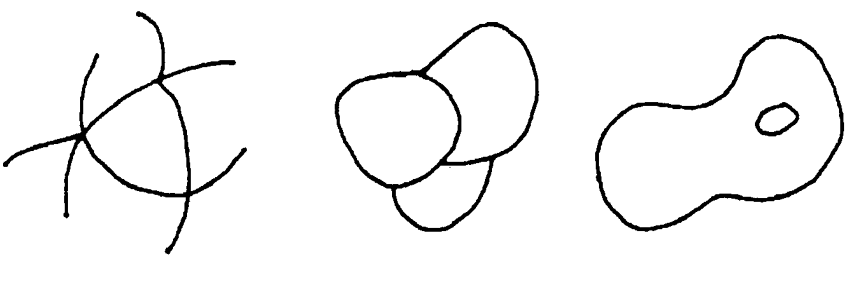


图 1-7：几种典型的拓扑关系

3）与几何位置无关的属性

即通常所说的非几何属性或简称属性，是与地理实体相联系的地理变量或地理意义。属性分为定性和定量的两种，前者包括名称、类型、特性等，后者包括数量和等级；定性描述的属性如土壤种类、行政区划等，定量的属性如面积、长度、土地等级、人口数量等。非几何属性一般是经过抽象的概念，通过分类、命名、量算、统计得到。任何地理实体至少有一个属性，而地理信息系统的分析、检索和表示主要是通过属性的操作运算实现的，因此，属性的分类系统、量算指标对系统的功能有较大的影响。

对地理信息系统的三种观点[Maguire,1991]

对于 GIS 的定义，有三种观点，即地图、数据库与空间分析的观点。地图观点的定义侧重于制图有关的内容，因此 GIS 被视为一个地图分析与处理系统；数据库观点的 GIS 定义则侧重于数据库设计与实现的完美性，一个复杂的数据库管理系统被视为 GIS 不可分割的一部分；持空间分析观点的定义侧重于分析和

---

* 在下文中，将不再区分“地理数据”和“空间数据”，它们都作为同一概念。

建模，GIS 被视为一门空间信息科学而不仅是一门技术。

3．地理信息系统的功能概述

地理信息系统的核心问题可归纳为五个方面的内容：位置、条件、变化趋势、模式和模型。

1）位置（Locations）

即在某个特定的位置有什么。

首先，必须定义某个物体或地区信息的具体位置，常用的定义方法有：通过各种交互手段确定位置，或者直接输入一个坐标；其次，指定了目标或区域的位置后，可以获得预期的结果以及其所有或部分特性，例如当前地块所有者、地址、土地利用情况、估价等。

2）条件（Conditions）

即什么地方有满足某些条件的东西。

首先，可以用下列方式指定一组条件，如从预定义的可选项中进行选取；填写逻辑表达式；在终端上交互地填写表格。

其次，指定条件后，可以获得满足指定条件的所有对象的列表，如在屏幕上以高亮度显示满足指定条件的所有特征，例如，其所位于的土地类型为居民区、估价低于 200,000 美元、有四个卧室而且是木制的的房屋。

3）变化趋势（Trends）

该类问题需要综合现有数据，以识别已经发生了或正在发生变化的地理现象。

首先，确定趋势，当然趋势的确定并不能保证每次都正确，一旦掌握了一个特定的数据集，要确定趋势可能要依赖假设条件、个人推测、观测现象或证据报道等。

其次，针对该趋势，可通过对数据的分析，对该趋势加以确认或否定。地理信息系统可使用户快速获得定量数据以及说明该趋势的附图等。例如，通过 GIS，可以识别该趋势的特性：有多少柑桔地块转作它用？现在作为何用？某一区域中有多少发生了这种变化？这种变化可回溯多少年？哪个时间段能最好反映该趋势？1 年、5 年还是 10 年？变化率是增加了还是减少了？

4）模式（Patterns）

该类问题是分析与已经发生或正在发生事件有关的因素。地理信息系统将现有数据组合在一起，能更好地说明正在发生什么，找出发生事件与哪些数据有关。

首先，确定模式，模式的确定通常需要长期的观察、熟悉现有数据、了解数据间的潜在关系。

其次，模式确定后，可获得一份报告，说明该事件发生在何时何地、显示事件发生的系列图件。例如，机动车辆事故常常符合特定模式，该模式（即事故）发生在何处？发生地点与时间有关吗？是不是在某种特定的交叉处？在这些交叉处又具有什么条件？

5）模型（Models）

该类问题的解决需要建立新的数据关系以产生解决方案。

首先，建立模型，如选择标准、检验方法等。

其次，建立了一个或多个模型后，能产生满足特定的所有特征的列表，并着重显示被选择特征的地图，而且提供一个有关所选择的特征详细描述的报表。例如要兴建一个儿童书店，用来选址的评价指标可能包括 10、15、20 分钟可到达的空间区域。附近居住的 10 岁或 10 岁以下的儿童的人数、附近家庭的收入情况、周围潜在竞争的情况。

为了完成上述的地理信息系统的核心任务，需要采用不同的功能来实现它们。尽管目前商用 GIS 软件包的优缺点是不同的，而且它们在实现这些功能所采用的技术也是不一样的，但是大多数商用 GIS 软件包都提供了如下功能：数据的获取（Data Acquisition）、数据的初步处理（Preliminary data Processing）、数据的存储及检索（Storage and Retrieval）、数据的查询与分析（Search and Analysis）、图形的显示与交互（Display and Interaction）。

图 1-8 说明了这些功能之间的关系，以及它们操作（Manipulation）数据的不同表现。

从图 1-8 中可以看出，数据获取是从现实世界的观测、以及从现存文件、地图中获取数据。有些数据

已经是数字化的形式，但是往往需要进行数据预处理，将原始数据转换为结构化的数据，以使其能够被系统查询和分析。查询分析是求取数据的子集或对其进行转换，并交互现实结果。在整个处理过程中，都需要数据存储检索以及交互表现的支持，换言之，这两项功能贯穿了地理信息系统数据处理的始终。

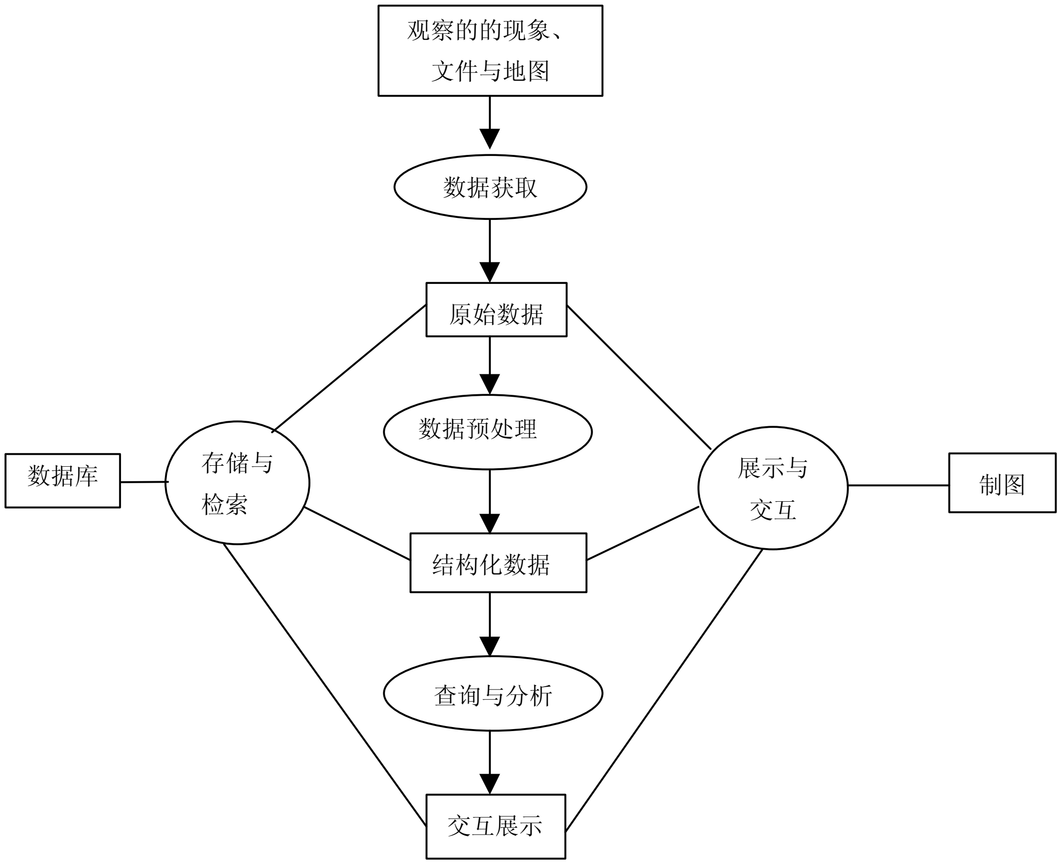


图 1-8：GIS 功能概述（椭圆）以及它们的表现（矩形）

空间信息处理和分析的六个组成部分[丁跃民]

1）空间操作，如地图的并、交、差运算，缓冲区计算，选择等等；

2）空间统计分析，用于描述和分析空间数据的关系，如空间自相关分析；

3）空间模型，注重于空间现象、空间结构、空间关系和空间位置的分析，如网络分析和水系生成等等；

4）空间表现/可视化，侧重于表达空间信息；

5）空间数据库管理，包括空间数据库设计，空间数据结构，空间数据管理和空间查询。

6）空间模型库管理，包括为空间决策支持系统提供模型的管理等等。

3．1 数据采集、监测与编辑

主要用于获取数据，保证地理信息系统数据库中的数据在内容与空间上的完整性、数值逻辑一致性与正确性等。一般而论，地理信息系统数据库的建设占整个系统建设投资的 70%或更多，并且这种比例在近期内不会有明显的改变。因此，信息共享与自动化数据输入成为地理信息系统研究的重要内容。目前可用于地理信息系统数据采集的方法与技术很多，有些仅用于地理信息系统，如手扶跟踪数字化仪；目前，自动化扫描输入与遥感数据集成最为人们所关注。扫描技术的应用与改进，实现扫描数据的自动化编辑与处理仍是地理信息系统数据获取研究的主要技术关键。

3．2 数据处理

初步的数据处理主要包括数据格式化、转换、概括。数据的格式化是指不同数据结构的数据间变换，是一种耗时、易错、需要大量计算量的工作，应尽可能避免；数据转换包括数据格式转化、数据比例尺的变化等。在数据格式的转换方式上，矢量到栅格的转换要比其逆运算快速、简单。数据比例尺的变换涉及到数据比例尺缩放、平移、旋转等方面，其中最为重要的是投影变换；制图综合（Generalization）包括数据平滑、特征集结等。目前地理信息系统所提供的数据概括功能极弱，与地图综合的要求还有很大差距，需要进一步发展。

3．3 数据存储与组织

这是建立地理信息系统数据库的关键步骤，涉及到空间数据和属性数据的组织。栅格模型、矢量模型或栅格/矢量混合模型是常用的空间数据组织方法。空间数据结构的选择在一定程度上决定了系统所能执行的数据与分析的功能；在地理数据组织与管理中，最为关键的是如何将空间数据与属性数据融合为一体。目前大多数系统都是将二者分开存储，通过公共项（一般定义为地物标识码）来连接。这种组织方式的缺点是数据的定义与数据操作相分离，无法有效记录地物在时间域上的变化属性。

3．4 空间查询与分析

空间查询是地理信息系统以及许多其它自动化地理数据处理系统应具备的最基本的分析功能；而空间分析是地理信息系统的核心功能，也是地理信息系统与其它计算机系统的根本区别，模型分析是在地理信息系统支持下，分析和解决现实世界中与空间相关的问题，它是地理信息系统应用深化的重要标志。地理信息系统的空间分析可分为三个不同的层次。

3．4．1 空间检索

包括从空间位置检索空间物体及其属性和从属性条件集检索空间物体。“空间索引”是空间检索的关键技术，如何有效地从大型的地理信息系统数据库中检索出所需信息，将影响地理信息系统的分析能力；另一方面，空间物体的图形表达也是空间检索的重要部分。

3．4．2 空间拓扑叠加分析

空间拓扑叠加实现了输入要素属性的合并（Union）以及要素属性在空间上的连接(Join)。空间拓扑叠加本质是空间意义上的布尔运算。

3．4．3 空间模型分析

在空间模型分析方面，目前多数研究工作着重于如何将地理信息系统与空间模型分析相结合。其研究可分三类：

第一类是地理信息系统外部的空间模型分析，将地理信息系统当作一个通用的空间数据库，而空间模型分析功能则借助于其它软件；

第二类是地理信息系统内部的空间模型分析，试图利用地理信息系统软件来提供空间分析模块以及发展适用于问题解决模型的宏语言，这种方法一般基于空间分析的复杂性与多样性，易于理解和应用，但由

于地理信息系统软件所能提供空间分析功能极为有限，这种紧密结合的空间模型分析方法在实际地理信息系统的设计中较少使用；

第三类是混合型的空间模型分析，其宗旨在于尽可能地利用地理信息系统所提供的功能，同时也充分发挥地理信息系统使用者的能动性。

3．5 图形与交互显示

地理信息系统为用户提供了许多用于地理数据表现的工具，其形式既可以是计算机屏幕显示，也可以是诸如报告、表格、地图等硬拷贝图件，尤其要强调的是地理信息系统的地图输出功能。一个好的地理信息系统应能提供一种良好的、交互式的制图环境，以供地理信息系统的使用者能够设计和制作出高质量的地图。

4．地理信息系统的研究内容

4．1 研究内容

地理信息系统是在地理学研究和生产实践的需求中产生，地理信息系统的应用使技术系统不断完善，并逐渐发展了地理信息系统的理论；理论研究又指导开发新一代高效地理信息系统，并不断拓宽其应用领域，加深应用的深度；地理信息系统的应用，又对理论研究和技术方法提出了更高的要求。这三个方面的研究内容是相互联系相互促进的。地理信息系统研究的内容主要有以下三个方面，见图 1-9 所示。

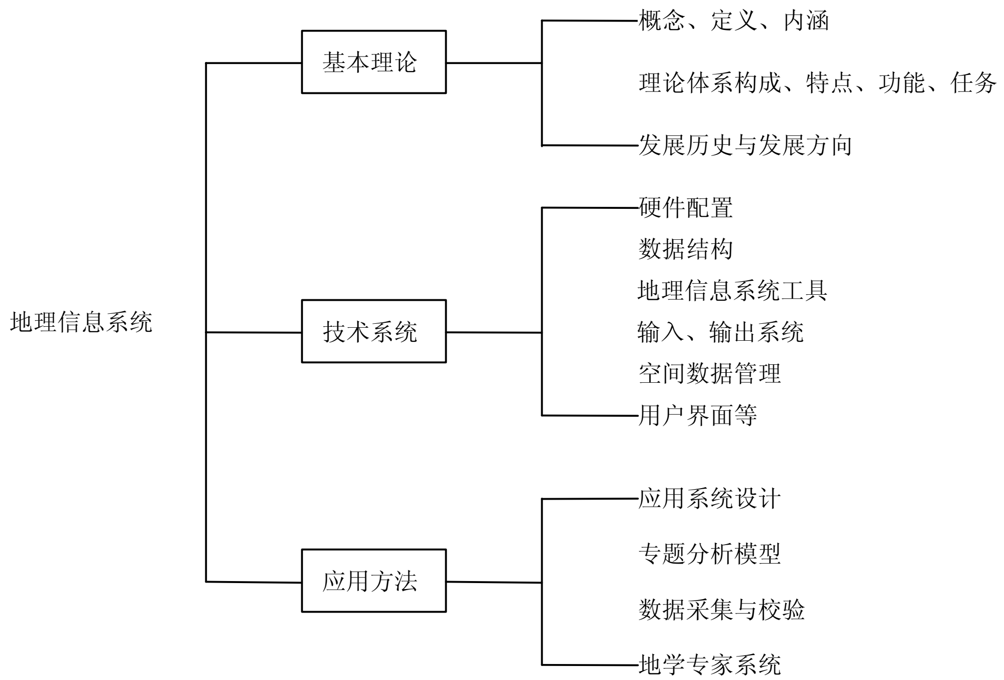


图 1-9：地理信息系统内容体系

4．1．1 地理信息系统基本理论研究

包括研究地理信息系统的概念、定义和内涵；地理信息系统的信息论研究；建立地理信息系统的理论

体系；研究地理信息系统的构成、功能、特点和任务；总结地理信息系统的发展历史，探讨地理信息系统发展方向等理论问题。

4．1．2 地理信息系统技术系统设计

包括地理信息系统硬件设计与配置；地理空间数据结构及表示；输入与输出系统；空间数据库管理系统；用户界面与用户工具设计；地理信息系统工具软件研制；微机地理信息系统的开发；网络地理信息系统的研制等。

4．1．3 地理信息系统应用方法研究

包括应用系统设计和实现方法；数据采集与校验；空间分析函数与专题分析模型；地理信息系统与遥感技术结合方法；地学专家系统研究等。

总之，地理信息系统的内容主要包括：有关的计算机软/硬件；空间数据的获取及计算机输入；空间数据模型及数字表达；数据的数据库存储及处理；数据的共享、分析与应用；数据的显示与视觉化；地理信息系统的网络化等等。

4．2 地理信息系统相关学科

地理信息系统，是六十年代开始迅速发展起来的地理学研究的新技术，是多种学科交叉的产物。作为传统科学与现代技术相结合的产物，地理信息系统为各种涉及空间数据分析的学科提供了新的方法，而这些学科的发展都不同程度地提供了一些构成地理信息系统的技术与方法。为了更好地掌握并深刻地理解地理信息系统，有必要认识和理解与地理信息系统相关的学科。

地理学是一门研究人类生活空间的学科，地理学研究空间分析的传统历史悠久，它为 GIS 提供了一些空间分析的方法与观点，成为 GIS 部分理论的依托。地理学的许多分支学科，如地图学、大地测量学等都与 GIS 有着密切的相依关系。另一方面，地理信息系统也以一种新的思想和新的技术手段解决地理学的问题，使地理学研究的数学传统得到充分发挥。地理信息系统的相关学科见图 1-10。

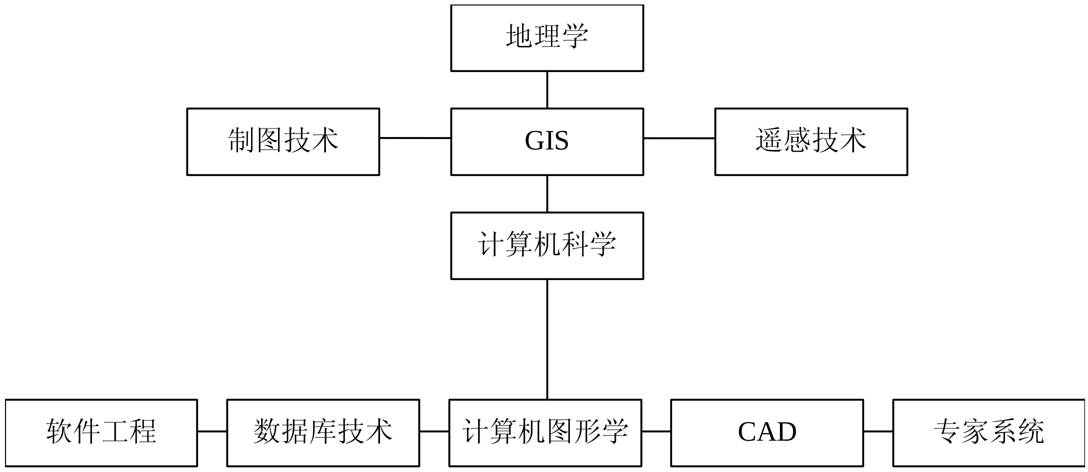


图 1-10：地理信息系统相关学科

4．2．1 地理学

地理学是一门研究人类赖以生存的空间的科学。在地理学研究中，空间分析的理论和方法具有悠久的

历史，它为地理信息系统提供了有关空间分析的基本观点与方法，成为地理信息系统的基础理论依托。而地理信息系统的发展也为地理问题的解决提供了全新的技术手段，并使地理学研究的数学传统得到了充分地发挥。

地理系统的内部及其外界，不仅存在着物质和能量的交流，还存在着信息流，这种信息交流使得系统许多似不相关的形态各异的要素联系起来，共同作用于地理系统。而地理信息系统体现着一种信息联系，由系统建立者输入，而由机器存储的各种影像、地图和图表都包含了丰富的地理空间信息的数据，通过指针或索引等组织信息相关联；系统软件对空间数据编码解码和处理；用户对 GIS 发出指令，GIS 按约定的方式做出解释后，获得用户指令信息，调用系统内的数据提取相应的信息，从而对用户做出反应，这是信息按一定方式流动的过程。

由此可见，地理信息系统不仅要以信息的形式表达自然界实体之间物质与能量的流动，更为重要的是以最直接的方式反映了自然界的信息联系，并可以快速模拟这种联系发展的结果，达到地理预测的目的。

总之，自然界与人类存在着深刻的信息联系，地理学家所面对的是一个形体的、即自然的地理世界，而感受到的却是一个地理信息世界。地理研究实际上是基于这个与真实世界并存而且在信息意义上等价的信息世界的，GIS 以地理信息世界表达地理现实世界，可以真实、快速地模拟各种自然的过程和思维的过程，对地理研究和预测具有十分重要的作用。

4．2．2 地图学

地图是记录地理信息的一种图形语言形式，从历史发展的角度来看，地理信息系统脱胎于地图，地图学理论与方法对地理信息系统的发展有着重要的影响。GIS 是地图信息的又一种新的载体形式，它具有存储、分析、显示和传输空间信息的功能，尤其是计算机制图为地图特征的数字表达、操作和显示提供了一系列方法，为地理信息系统的图形输出设计提供了技术支持；同时，地图仍是目前地理信息系统的重要数据来源之一。但二者又有本质之区别：地图强调的是数据分析、符号化与显示，而地理信息系统更注重于信息分析。

地图是认识和分析研究客观世界的常用手段，尽管地图的表现形式发生了种种变化，但是依然可以认为构成地图的主要因素有三：地图图形、数学要素和辅助要素。地图图形是用地图符号所表示的制图区域内，各种自然和社会经济现象的分布、联系以及时间变化等的内容部分（又称地理要素），如江河山地、平原、土质植被、居民点、道路、行政界限或其他专题内容等，这是地图构成要素中的主体部分。数学要素是决定图形分布位置和几何精度的数学基础，是地图的“骨架”。其中包括地图投影及坐标网、比例尺、大地控制点等。地图投影是用数学方法将地球椭球面上的图形转绘到平面上；坐标网是各种地图的数学基础，是地图上不可缺少的要素；比例尺表示坐标网和地图图形的缩小程度；大地控制点是保证将地球的自然表面转绘到椭球面上，再转绘到平面直角坐标网内时，具有精确的地理位置。辅助要素是为了便于读图与用图而设置的。如图例就是显示地图内容的各种符号的说明，还有图名、地图编制和出版单位、编图时间和所用编图资料的情况、出版年月等。有的地图上还有补充资料，用以补充和丰富地图的内容。如在图边或图廓内空白处，绘制一些补充地图或剖面图、统计图等。有时还有一些表格或某一方面的重点文字说明。

从地理信息系统的发展过程可以看出，地理信息系统的产生、发展与制图系统存在着密切的联系，两者的相通之处是基于空间数据库的表达、显示和处理。从系统构成与功能上看，一个地理信息系统具有机助制图系统的所有组成和功能，并且地理信息系统还有数据处理的功能。地图是一种图解图像，是根据地理思想对现实世界进行科学抽象和符号表示的一种地理模型，是地理思维的产物，也是实体世界地理信息的高效载体，地图可以从不同方面、不同专题，系统地记录和传输实体世界历史的、现在的和规划预测的地理景观信息。

4．2．3 计算机科学

地理信息系统技术的创立和发展是与地理空间信息的表达、处理、分析和应用手段的不断发展分不开的。二十世纪六十年代初，在计算机图形学的基础上出现了计算机化的数字地图。地理信息系统与计算机的数据库技术(DBMS)、计算机辅助设计(CAD)、计算机辅助制图(CAM)和计算机图形学(Computer Graphics)等有着密切的联系。但是它们却无法取代地理信息系统的作用。

数据库管理系统(Database Management System)是操作和管理数据库的软件系统，提供可被多个应用程序和用户调用的软件系统，支持可被多个应用程序和用户调用的数据库的建立、更新、查询和维护功能，GIS 在数据管理上借鉴 DBMS 的理论和方法，非几何属性数据有时也采用通用 DBMS 或在其上开发的软件系统管理；对于空间地理数据的管理，通用的 DBMS 有两个明显的弱点：第一，缺乏空间实体定义能力：目前流行的网状结构、层次结构、关系结构等，都难以对空间结构全面、灵活、高效地加以描述；第二，缺乏空间关系查询能力：通用的 DBMS 的查询主要是针对实体的查询，而 GIS 中则要求对实体的空间关系进行查询，如关于方位、距离、包容、相邻、相交和空间覆盖关系等，显然，通用 DBMS 难以实现对地理数据空间查询和空间分析。数据是信息的载体，对数据进行解释可提取信息，通用数据库和地理数据库都是针对数据本身进行管理，而 GIS 则在数据管理基础上，通过地理模型运算，产生有用的地理信息，取得信息的多少和质量，与地理模型的水平密切相关。

计算机图形学是利用计算机处理图形信息以及借助图形信息进行人——机通讯处理的技术，是 GIS 算法设计的基础。GIS 是随着计算机图形学技术的发展而不断发展完善的，但是计算机图形学所处理的图形数据是不包含地理属性的纯几何图形，是地理空间数据的几何抽象，可以实现 GIS 底层的图形操作，但不能完成数据的地理模型分析和许多具有地理意义的数据处理，不能构成完整的 GIS。

计算机辅助设计(CAD,Computer-Aided Design)是通过计算机辅助设计人员进行设计，以提高设计的自动化程度，节省人力和时间；专门用于制图的计算机辅助制图(Computer-Aided Mapping)，采用计算机进行几何图形的编辑和绘制。GIS 与 CAD 和 CAM 的区别在于：第一，CAD 不能建立地理坐标系和完成地理坐标变换；第二，GIS 的数据量比 CAD、CAM 大得多，结构更为复杂，数据间联系紧密，这是因为 GIS 涉及的区域广泛，精度要求高，变化复杂，要素众多，相互关联，单一结构难以完整描述；第三，CAD 和 CAM 不具备 GIS 具有地理意义的空间查询和分析功能。

4．2．4 遥感

遥感是一种不通过直接接触目标物而获得其信息的一种新型的探测技术。它通常是指获取和处理地球表面的信息，尤其是自然资源与人文环境方面的信息，并最后反映在像片或数字影像上的技术。影像通常需要进一步处理方可使用，用于该目的的技术称为图像处理。图像处理包括各种可以对像片或数字影像进行处理的操作，这些操作包括影像压缩，影像存储，影像增强、处理以及量化影像模式识别等。目前，遥感已经成为环境研究中极有价值的工具，不同学科的专业人员不断地发现航空遥感不同数据在各领域内的潜在应用。遥感和图像处理技术被用于获取和处理地球表面有关的信息；GIS 的发展则源于对土地属性信息与相应几何表达的集成及空间分析的需求。这两项技术在过去是相互独立发展的，尽管他们实际上是互补的。从地理信息系统本身的角度出发，随着它应用领域的开拓和深入，它首先要求存储大量的有关数据，通过不断的积累和延伸，从而具备反映自然历史过程和人为影响的趋势的能力，揭示事物发展的内在规律。但是地理信息系统数据库几乎只是通过地图数字化建立起来的，用户不能接触到原始资料及其有关信息，而地理信息系统中的原始数据却是有效地模拟和控制误差传播的基础。其次，地理信息系统为了保持系统的动态性和现势性，它还要求及时地更新系统中的数据，目前地理信息系统中存储的信息只是现实世界的一个静态模型，需要定时或及时的更新。遥感作为一种获取和更新空间数据的强有力手段，能及时地提供准确、综合和大范围内进行动态检测的各种资源与环境数据，因此遥感信息就成为地理信息系统十分重要的信息源。另一方面 GIS 中的数据可以作为遥感影像分析的一种辅助数据。在两者集成过程中，GIS 主要用于数据处理、操作和分析；而遥感则作为一种数据获取、维护与更新 GIS 中的数据的手段，此外，GIS

可用于基于知识的遥感影像分析。地理信息系统和遥感是两个相互独立发展起来的技术领域，随着它们应用领域的不断开拓和自身的不断发展，即由定性到定量、由静态到动态、由现状描述到预测预报的不断深入和提高。它们的结合也逐渐由低级向高级阶段发展。遥感和地理信息系统的结合经历了由低级向高级阶段的发展过程。最早的结合工作包括把航空遥感像片经目视判读和处理后编制成各种类型的专题图，然后将它们数字化和输入地理信息系统；从 70 年代中后期开始，各种影像分析系统得到了迅速而广泛地发展。大量的遥感数据以及图像分析系统图像分类所形成的各类专题信息，可以直接输入地理信息系统，整个过程能在“全数字”的环境下进行，图像数据能够在生成编辑地图的屏幕上显示，标志着遥感和地理信息系统的结合进入了新的阶段。

遥感作为空间数据采集手段，已成为地理信息系统的主要信息源与数据更新途径。遥感图像处理系统包含若干复杂的解析函数，并有许多方法用于信息的增强与分类。另外，大地测量为地理信息系统提供了精确定位的控制系统，尤其是全球定位系统（GPS）可快速、廉价地获得地表特征的熟悉位置信息。航空像片及其精确测量方法的应用使得摄影测量成为地理信息系统主要的地形数据来源。总之，遥感是地理信息系统的主要数据源与更新手段，同时，地理信息系统的应用又进一步支持遥感信息的综合开发与利用。

4．2．5 管理科学

传统意义上的管理信息系统是以管理为目的，在计算机硬件和软件支持下具有存储、处理、管理和分析数据能力的信息系统，如人才管理信息系统、财务管理信息系统、服务业管理信息系统等。这类信息系统的最大特征是它处理的数据没有或者不包括空间特征。

另一类管理信息系统是以具有空间分析功能的地理信息系统为支持、以管理为目标的信息系统，它利用地理信息系统的各种功能实现对具有空间特征的要素进行处理分析以达到管理区域系统的目的，如城市交通管理信息系统、城市供水管理信息系统、节水农业管理信息系统等。

事实上，可以形象地把地理信息系统与其他学科的关系用一棵树来表示，如图 1-11 所示。

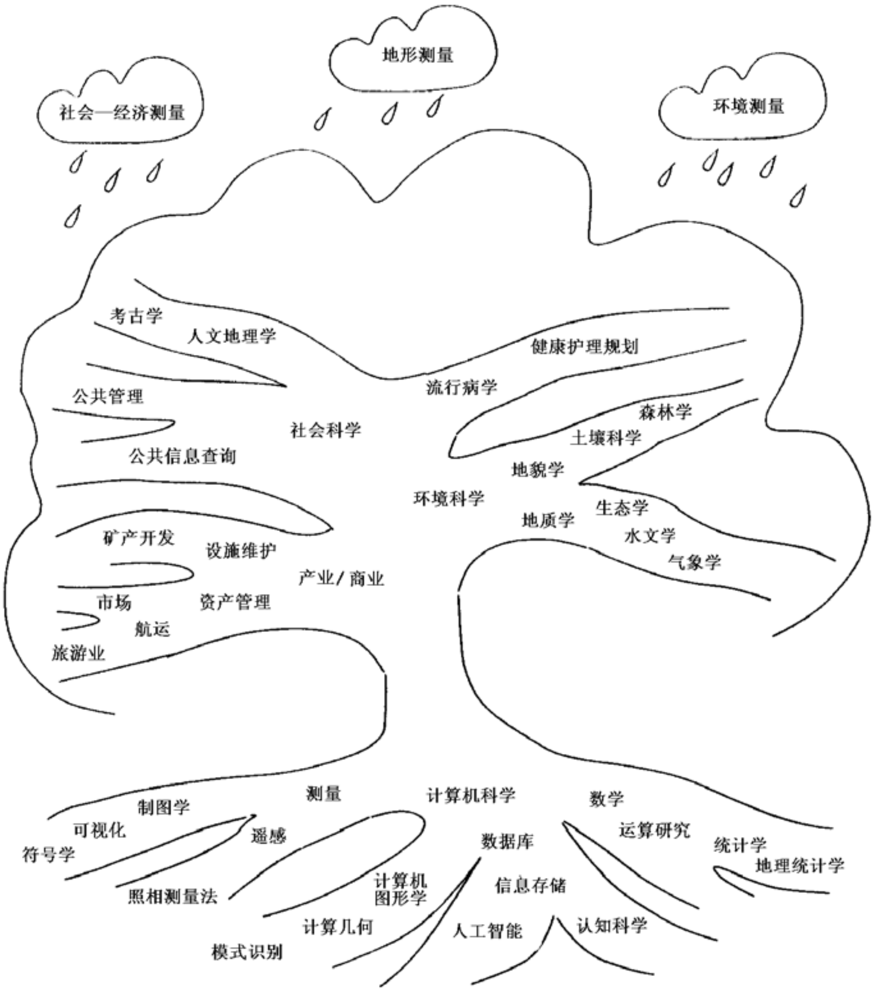


图 1-11：GIS 学科“树” [Charistopher.B.Jones]

正如上图所述，“树根”表示 GIS 的技术基础，如测量学、计算机科学与数学等；“树枝”表示 GIS 的应用，应用的结果与需求返回到“树根”；“雨滴”是每个应用中的数据来源，如各种测量如地形测量、环境测量等，并为它的发展提供了有效的手段，而地理信息系统的应用主要是在环境科学、地理学和社会科学等领域。

5．地理信息系统发展简史

地理信息系统脱胎于地图，它们都是地理信息的载体，具有获得、存储、编辑、处理、分析与显示地理数据的功能。地图是地理学的第二代语言，而地理信息系统将成为地理学的第三代语言。二十世纪六十年代初，在计算机图形学的基础上出现了计算机化的数字地图。1950 年，麻省理工学院为它的旋风一号计算机制造了第一台图形显示器；1958 年，美国的一公司在联机的数字记录仪的基础上研制成滚筒式绘图仪；1962 年，麻省理工学院的一名研究生在其博士学位论文中，首次提出了计算机图形学的术语，并论证了交互式计算机图形学是一个可行的、有用的研究领域，从而确立了这一科学分支的独立地位。在此基础上，地理信息系统发展起来。

5．1 六十年代开拓发展阶段

六十年代初，计算机技术开始用于地图量算、分析和制作，由于机助制图具有快速、廉价、灵活多样、

易于更新、操作简便、质量可靠、便于存储、量测、分类、合并和覆盖分析等优点而迅速发展起来。六十年代中期，由于对于自然资源和环境的规划管理和应用加速增长的需要，对大量空间环境数据存储、分析和显示技术方法改进的要求，以及计算机技术及其在自然资源和环境数据处理中应用的迅速发展，促使对地图进行综合分析和输出的系统日益增多。

六十年代中后期，许多与 GIS 有关的组织和机构纷纷建立并开展工作，如美国城市和区域系统协会（URISA）在 1966 年成立，美国州信息系统全国协会（NASIS）在 1969 年成立，城市信息系统跨机构委员会（UAAC）在 1968 年成立，国际地理联合会（IGU）的地理数据遥感和处理小组委员会在 1968 年成立等。这些组织和机构相继组织了一系列地理信息系统的国际讨论会。

最初的系统主要是关于城市和土地利用的，如加拿大地理信息系统（CGIS）就是为处理加拿大土地调查获得的大量数据建立的。该系统由加拿大政府组织于 1963 年开始研制实施，到 1971 年投入正式运行，被认为是国际上最早建立的、较为完善的大型使用的地理信息系统。

由于计算机硬件系统功能较弱，限制了软件技术的发展。这一时期地理信息系统软件的研制主要是针对具体的 GIS 应用进行的，到六十年代末期，针对 GIS 一些具体功能的软件技术有了较大进展。

第一，栅格——矢量转换技术、自动拓扑编码以及多边形中拓扑误差检测等方法得以发展，开辟了分别处理图形和属性数据的途径；

第二，具属性数据的单张或部分图幅可以与其他图幅或部分在图边自动拼接，从而构成一幅更大的图件，使小型计算机能够分块处理较大空间范围（或图幅）的数据文件；

第三，采用命令语言建立空间数据管理系统，对属性再分类、分解线段、合并多边形、改变比例尺、测量面积、产生图和新的多边形、按属性搜索、输出表格和报告以及多边形的叠加处理等。

这一时期的软件主要针对当时的主机和外设开发的，算法尚嫌粗糙，图形功能有限。

5．2 七十年代巩固阶段

进入七十年代以后，由于计算机硬件和软件技术的飞速发展，尤其是大容量存取设备——硬盘的使用，为空间数据的录入、存储、检索和输出提供了强有力的手段。用户屏幕和图形、图像卡的发展增强了人机对话和高质量图形显示功能，促使 GIS 朝着使用方向迅速发展。一些发达国家先后建立了许多不同专题、不同规模、不同类型的各具特色的地理信息系统。如美国森林调查局发展了全国林业统一使用的资源信息显示系统；美国地质调查所发展了多个地理信息系统用于获取和处理地质、地理、地形和水资源信息较典型的有 GIRAS；日本国土地理院从 1974 年开始建立数字国土信息系统，存储、处理和检索测量数据、航空像片信息、行政区划、土地利用、地形地质等信息，为国家和地区土地规划服务；瑞典在中央、区域和市三级上建立了许多信息系统，比较典型的如区域统计数据库、道路数据库、土地测量信息系统、斯德哥尔摩地理信息系统、城市规划信息系统等；法国建立了地理数据库 GITAN 系统和深部地球物理信息系统等。

此外，探讨以遥感数据为基础的地理信息系统逐渐受到重视，如将遥感纳入地理信息系统的可能性、接口问题以及遥感支持的信息系统的结构和构成等问题；美国喷气推动实验室（JPL）在 1976 年研制成功兼具影像数据处理和地理信息系统功能的影像信息系统 IBIS（Image Based Information System），可以处理 Landsat 影像多光谱数据；NASA 的地球资源实验室在 1979 年至 1980 年发展了一个名为 ELAS 的地理信息系统，该系统可以接受 Landsat MSS 影像数据、数字化地图数据、机载热红外多波段扫描仪以及海洋卫星合成孔径雷达的数据等，产生地面覆盖专题图。

由于这一时期 GIS 的需求增加，许多团体、机构和公司开展了 GIS 的研制工作，推动 GIS 软件的发展。据 IGU 地理数据遥测和处理小组委员会 1976 年的调查，处理空间数据的软件已有 600 多个，完整的 GIS 有 80 多个。这一时期地图数字化输入技术有了一定的进展，采用人机交互方式，易于编辑修改，提高了工作效率，扫描输入技术系统出现。图形功能扩展不大，数据管理能力也较小。这一时期软件最重要的进展是人机图形交互技术的发展。

5．3 八十年代突破阶段

由于计算机的发展，推出了图形工作站和个人计算机等性能价格比大为提高的新一代计算机，计算机和空间信息系统在许多部门广泛应用。随着计算机软、硬件技术的发展和普及，地理信息系统也逐渐走向成熟。这一时期是地理信息系统发展的重要时期。计算机价格的大幅度下降，功能较强的微型计算机系统的普及和图形输入、输出和存储设备的快速发展，大大推动了地理信息系统软件的发展，并研制了大量的微机 GIS 软件系统。由于微机系统的软件环境限制较严，使得在微机 GIS 中发展的许多算法和软件技术具有很高的效率，GIS 软件技术在以下几个方面有了很大的突破。在栅格扫描输入的数据处理方面，尽管扫描数据的处理要花费很长的机时（与扫描时间相比为 10：1），但是仍可大大提高数据输入的效率；在数据存储和运算方面，随着硬件技术的发展，GIS 软件处理的数据量和复杂程度大大提高，许多软件技术固化到专用的处理器中；而且遥感影像的自动校正、实体识别、影像增强和专家系统分析软件也明显增加；在数据输出方面，与硬件技术相配合，GIS 软件可支持多种形式的地图输出；在地理信息管理方面，除了 DBMS 技术已发展到支持大型地图数据库的水平外，专门研制的适合 GIS 空间关系表达和分析的空间数据库管理系统也有了很大的发展。

总之，这一时期的地理信息系统的发展有如下特点：

第一，在七十年代技术开发的基础上，地理信息系统技术全面推向应用；

第二，开展工作的国家和地区更为广泛，国际合作日益加强，开始探讨建立国际性的地理信息系统，地理信息系统由发达国家推向发展中国家，如中国；

第三，地理信息系统技术进入多种学科领域，从比较简单的、单一功能的、分散的系统发展到多功能的、共享的综合性信息系统，并向智能化发展，新型的地理信息系统将运用专家系统知识，进行分析、预报和决策；

第四，微机地理信息系统蓬勃发展，并得到广泛应用。在地理信息系统理论指导下研制的地理信息系统工具具有高效率和更强的独立性和通用性，更少依赖于应用领域和计算机硬件环境，为地理信息系统的建立和应用开辟了新的途径。

我国地理信息系统方面的工作自八十年代初开始。以 1980 年中国科学院遥感应用研究所成立的全国第一个地理信息系统研究室为标志，在几年的起步发展阶段中，我国地理信息系统在理论探索、硬件配制、软件研制、规范制定、局部系统建立、初步应用实验和技术队伍培养等方面都取得了进步，积累了经验，为全国范围内开展地理信息系统的研制和应用奠定了基础。

5．4 九十年代社会化阶段

进入九十年代，随着地理信息产业的建立和数字化信息产品在全世界的普及，地理信息系统将深入到各行各业乃至各家各户，成为人们生产、生活、学习和工作中不可缺少的工具和助手。地理信息系统已成为许多机构必备的工作系统，尤其是政府决策部门在一定程度上由于受地理信息系统影响而改变了现有机构的运行方式、设置与工作计划等。而且，社会对地理信息系统认识普遍提高，需求大幅度增加，从而导致地理信息系统应用的扩大与深化。国家级乃至全球性的地理信息系统已成为公众关注的问题。

自九十年代起，中国地理信息系统步入快速发展阶段。力图使地理信息系统从初步发展时期的实验、局部应用走向实用化和生产化，为国民经济重大问题提供分析和决策依据。同时地理信息系统的研究和应用正逐步形成行业，具备了走向产业化的条件。

表 1-1：从 60 年代以来地理信息系统发展中重要的历史事件

1960	美国空军 CIA 首次成功地发射 CORONA
1963	Roger Tomlinson 开始了加拿大地理信息系统的开发
1963	Dr. Edgar Horwood 建立了城市与区域信息系统联合会(URISA)
1964	Howard Fisher 建立了计算机图形和空间分析的哈佛实验室
1966	SYMAP 系统在西北技术学院研制并在哈佛实验室完成

1967	DIME（双重独立制图编码）为美国人口普查局所研制
1969	Jack 和 Laura Dangermond 建立了环境系统研究所(ESRI)
1969	Jim Meadlock 建立了 Integraph 公司
1969	在英国诞生了激光扫描仪
1969	Ian McHarg 很有影响的书“自然设计(Design With Nature)”出版
1971	加拿大地理信息（CGIS）建立
1972	IBM 的 GFIS 发布
1972	GISP（General Information System for Planning）开发
1972	Landsat 卫星首次发射成功
1973	USGS 研制了地理信息提取和分析系统
1973	马里兰自动地理信息(MAGI, Maryland Automatic Geographic Information)开发
1974	在伦敦的皇家艺术学院建立了试验制图单元(ECU, Experimental Cartography Unit)
1974	首次自动制图会议在 Reston，弗吉尼亚召开
1976	明尼苏达研制了明尼苏达土地管理信息系统
1977	USGS 研制了数字化线图(DLG)空间数据模式
1978	ERDAS 成立
1978	地图叠加复合与统计系统开发
1979	哈佛图形实验室研制了 ODYSSEY GIS
1981	ESRI ARC/ INFO GIS 发布
1982	NASA 发射了 Landsat TM4
1983	ETAK 数字制图公司成立
1984	Marble, Calkins & Peuquet 出版了“地理信息系统的基本读物”（Basic Readings in Geographic Information Systems）
1984	第一届国际空间数据处理会议召开
1984	Landsat 商业化
1984	NASA 发射 Landsat TM5
1985	GPS 成为可运行系统
1985	美国军队建筑工程实验室开始研制 GRASS（Geographic Resources Analysis Support Systems，地理资源分析支持系统）
1986	MapInfo 建立
1986	Peter Burrough 出版了“土地资源评估的地理信息系统原理”(Principles of Geographic Information Systems for Land Resources Assessment)
1986	SPOT 卫星首次发射
1987	“地理信息系统的国际杂志” 出版
1987	Tydac SPANS GIS 发布
1987	科拉克大学开始 Idrisi 项目
1988	美国人口调查局第一次公开发布 TIGER
1988	纽约州立大学开始研制 GIS-L Internet list-server
1988	GIS World 首次发行
1988	首次 GIS/ LIS 会议举行
1988	英国的区域研究实验室成立
1988	Small World 公司成立
1989	在英国成立了地理信息系统联合会(AGI)
1989	Stan Arnoff 出版了“地理信息系统：一个管理透视” （Geographic Information Systems: a Management Perspective）
1989	Intergraph 发布 MGE
1991	Maguire, Goodchild 和 Rhind 出版了“地理信息系统：原理和应用”
1992	MAPS ALIVE 发行
1993	Digital Matrix Systems 发布了 InFoCAD for Windows NT 第一个版本，它是第一个基于 Win NT 的 GIS 软件
1994	OGC 形成（David Schell, Ken Gardells, Kurt Buehler,et al）
1995	MapInfo 专业版发布
1999	NASA 发射了 Landsat TM7

6．GIS 的发展展望

6．1GIS 理论研究中及待解决的问题

6．1．1GIS 理论发展的需求

GIS 是一门技术引导的多技术交叉的信息空间科学，它是对地理信息数据（包括图形和非图形数据，几何数据与属性数据）进行采集、存储、加工和再现，并能回答一系列问题的计算机系统，所以它必然是技术导向的。GIS 不断地用新的技术和方法来装备和发展自己，它在技术上所关注的是：数据采集；数据建模；数据的精度和系统回答问题的可信度；数据量；数据存取与保密；数据分析；用户接口；成本与效益；GIS 系统的寿命；GIS 系统工作的组织问题。这些技术问题，将会随着相关学科和软件、硬件手段的不断进步，而日趋完善。同时，GIS 是一门以应用为目的的信息产业，即 GIS 也是应用导向的，即它除了具有基础性和公益性特点，服务于科学研究和造福人类外，它还具有实际应用并创造价值的广阔市场。GIS 的应用可以深入到各个领域、各个机构，形成诸如资源 GIS，灾害监测和防治 GIS，农林牧副渔 GIS 等。GIS 的不断发展，它既依赖于地理学、统计学和测量学这些基础学科，又取决于计算机软件技术、航天技术、遥感技术和人工智能与专家系统技术的进步与成就。它是位于地学与技术科学的边缘，但本质上已是信息科学的一个组成部分。

随着 GIS 理论的发展与完善，以及人们对空间信息需求量的增大，为了使得 GIS 系统得到可持续的发展，则必须使 GIS 向集成化和智能化方向发展，该方向包括以下几个方面：

1）图形数据和属性数据的结合：最初，图形和属性数据是完全分开的；然后，通过内部连接，将二者联系起来；再次，进行了混合处理；最后，达到完全的结合。未来 GIS 要求的是将二者的完全的结合。

2）GIS 与 RS 的结合：遥感是地理信息系统重要的数据源和数据更新的手段。相反，GIS 则是遥感中数据处理的辅助信息，用于语义和非语义信息的自动提取。GIS 与 RS 可能的结合方式包括：分开但是平行的结合（不同的用户界面、不同的工具库和不同的数据库）、表面无缝的结合（同一用户界面，不同的工具库和不同的数据库）和整体的结合（同一个用户界面、工具库和数据库）。未来要求的是整体的结合。

3）GIS 与 GPS 和 CCD 技术的结合：GPS 是全球定位系统，利用 GPS 接收机，可以直接测定地面上任一点的三维坐标。GPS 与 GIS 相结合可以实现电子导航，用于交通管理、公安侦破、自动导航，也可以用作 GIS 实时更新。如果再加上 CCD 摄象机实时摄象和配以影像处理，则可以形成实时 GIS 运行系统，用于公路、铁路线路状况的自动监测和管理，以及作战指挥系统等。

4）GIS 与专家系统（Expert System）的结合：由于 GIS 是一个基于地理数据的空间信息系统，它必须具有自动采集和处理数据的功能，而且能够智能化地分析和运用数据，提供科学的决策咨询，以回答用户可能提出的各种复杂问题。从这个意义 GIS 与 ES 相结合，形成智能化的高度集成 GIS 系统。

6．1．2 GIS 理论研究中及待解决的问题

基于以上的分析，要想得到一个理想的 GIS 系统，需要 GIS 理论上解决以下的主要问题：

1）GIS 设计与实现的方法学问题

由于缺乏严格的工程管理和好的分析设计方法支持，导致了 GIS 软件系统的可靠性和可维护性差。这是一个长期以来人们一直在尽力解决但还未解决的问题。

2）GIS 的功能问题

当前以数据采集、存储、管理和查询检索功能为主的 GIS，还不能完全满足社会和区域可持续发展在空间分析、预测预报、决策支持等方面的要求，直接影响到 GIS 的应用效益和生命力。

3）多媒体地理信息系统的管理和操作的问题

在一个多种数据类型并存的混合系统中，如何实现对各类数据的随意操作和有效管理，这是现今信息媒体多元化新时代的突出问题，它比单一地图数据库的操作和管理更复杂。

4）GIS 地理信息的深加工问题

目前的 GIS 还远未发挥它提供结论性专题地图和数据集方面的作用，这是涉及对 GIS 地理信息进行深加工的问题。这种深加工的结果，可以是结论性专题地图，也可以是结论性专题数据集。这两种形式都是必须的，前者提供结论性图形信息，后者提供结论性数字信息，提供经过深加工的结论性成果对用户更直接和更有利。

5）空间信息可视化技术和虚拟现实技术（VR）

可视化技术已经远远超过了传统的符号化及视觉变量表示法的水平，进入了在动态、时空变换、多维的可交互的地图条件下探索视觉效果和提高视觉工具功能的阶段，它的重点是要将那些通常难于设想和接近的环境与事物，以动态直观的方式表现出来。GIS 可视化方面的研究主要集中在以下几个方面：运用动画技术制作动态地图，可用于涉及时空变化的现象或概念的可视化分析；运用 VR 技术进行地形环境仿真，真实再现地景，用于交互式地观察和分析，提高对地形环境的认知效果；运用图形显示技术进行空间数据的不确定性和可靠性的检查，把抽象数据可视化，由此发现规律；运用图形界面和交互式手段进行地图设计和编辑，以直观的方式完成地图设计制作（如地图颜色的可视化设计）；可视化技术用于视觉感受及空间认知理论的研究。

6．2 地理信息系统的发展动态

近年来地理信息系统技术发展迅速，其主要的原动力来自日益广泛的应用领域对地理信息系统不断提出的要求。另一方面，计算机科学的飞速发展为地理信息系统提供了先进的工具和手段，许多计算机领域的新技术，如面向对象技术、三维技术、图像处理和人工智能技术都可直接应用到地理信息系统中。下面对当前地理信息系统研究中的几个热点研究领域作一介绍。

6．2．1GIS 软件中发展的热点

1）GIS 中面向对象技术研究

面向对象方法为人们通过计算机直接描述现实世界提供了一条适合于人类思维模式的方法，面向对象的技术在 GIS 中的应用，即面向对象的 GIS，已成为 GIS 的发展方向。这是因为空间信息较之传统数据库处理的一维信息更为复杂、繁琐，面向对象的方法为描述复杂的空间信息提供了一条直观、结构清晰、组织有序的方法，因而倍受重视。

面向对象的 GIS 较之传统的 GIS 有下列优点：

1）所有的地物以对象形式封装，而不是以复杂的关系形式存储，使系统组织结构良好、清晰；

2）以对象为基础，消除了分层的概念；

3）面向对象的一般——特殊结构和整体——部分结构使 GIS 可以直接定义和处理复杂的地物类型；

4）根据面向对象后期绑定*（Late-binding）的思想，用户可以在现有抽象数据类型和空间操作箱上定义自己所需的数据类型和空间操作方法，增强系统的开发性和可扩充性；

5）基于图标的面向对象的用户界面，便于用户操作和使用。面向对象的 GIS 也存在一些尚待进一步研究的问题：

1）大对象的操作仍受硬件条件的限制；

2）对象的独立性与粒度问题；

3）矢量和栅格数据统一的、支持动态拓扑结构和复合对象表示的面向对象的数据结构问题。

---

* 后期绑定是面向对象编译器中的重要概念，采用该技术，一个对象所属的类可以在运行时刻（Run Time）指定，而不是在编译成目标码时确定。

2）时空系统

传统的地理信息系统只考虑地物的空间特性，忽略了其时间特性。在许多应用领域中，如环境检测、地震救援、天气预报等，空间对象是随时间变化的，而这种动态变化规律在求解过程中起着十分重要的作用。过去 GIS 忽略时态主要是受软硬件条件的限制，也有技术方面的原因。近年来，对 GIS 中时态特性的研究变得十分活跃，即所谓"时空系统"。通常把 GIS 的时间维分成处理时间维和有效时间维，处理时间又称数据库时间或系统时间，它指在 GIS 中处理发生的时间；有效时间亦称事件时间或实际时间，它指在实际应用领域事件出现的时间。时空系统主要研究时空模型，时空数据的表示、存储、操作、查询和时空分析。目前比较流行的作法是在现有数据模型基础上扩充，如在关系模型的元组中加入时间，在对象模型中引入时间属性。在这种扩充的基础上如何解决从表示到分析的一系列问题仍有待进一步研究。

3）地理信息建模系统（GIMS-Geographic Information Modeling System）

通用 GIS 的空间分析功能对于大多数的应用问题是远远不够的，因为这些领域都有自己独特的专用模型，目前通用的 GIS 大多数通过提供二次开发的工具和环境来解决这一问题，但二次开发工具的一个主要问题是它对于普通用户而言过于困难。而 GIS 成功应用到专门领域的关键在于支持建立该领域特有的空间分析模型。GIMS 的研究动向为：面向对象在 GIS 中的应用。面向对象技术用对象（实体属性和操作的封装）、对象类结构（分类和组装结构）、对象间的通讯来描述客观世界，为描述复杂的三维空间提供了一条结构化的途径；基于图标（Icon）的用户建模界面。建模过程中的对象和空间分析操作均以图标形式展示给用户，用户亦可以自定义图标。用户对图标的定义、选择和操作中完成模型的定义和检验。这种方法较之 AML 这类宏语言要方便和直观得多；GIS 与其他的模型和知识库的结合，这是许多应用领域面临的一个非常实际的问题，即存在 GIS 之外的模型和知识库如何与 GIS 耦合成一个有机整体。

4）三维地理信息系统的研究

三维 GIS 是许多应用领域对 GIS 的基本要求。目前的 GIS 大多数提供了一些较为简单的三维显示和操作功能，但这与真三维表示和分析还有很大差距。真正的三维 GIS 必须支持真三维的矢量和栅格数据模型及以此为基础的三维空间数据库，解决了三维空间操作和分析问题。主要研究的方向包括：三维数据结构的研究，主要包括数据的有效存储、数据状态的表示和数据的可视化；三维数据的生成和管理；地理数据的三维显示，主要包括三维数据的操作，表面处理，栅格图像、全息图像显示，层次处理等。

6．2．2 实用地理信息系统发展趋势与展望

随着计算机和信息技术的发展，GIS 迅速地变化着。在未来 10 年内的 GIS 发展用下面几个方面来概括。

1）GIS 网络化

对于 GIS 的发展，计算机网络技术是起到质变作用的重要技术。它的发展使得以往很多难以完成的事情得以实现，如网络技术使得数据库在地理位置上以分布的方式存在，这样做，各个数据库可以局部地进行生产、更新、维护和管理，而网络又使这些分布在局部的数据库相互之间可以连接起来实现共享使用。高速度的数据传输使得数据库之间的数据传输能够快速地实现。万维网的发展给 GIS 数据在更大范围内的发布、出版、获取和查询提供了有效可行的途径。网络浏览器的使用从视觉上给提供和使用地理数据的人们带来了方便。地理数据不仅可以按照地理位置、专题内容、生产机构、使用价格等进行搜索，甚至可以直接在网上进行数据的各类空间操作，使用网络提供的各类模型进行模拟，直接产生新的数据结果，真正地实现"网络就是计算机"这一新的概念模式。网络技术虽然发展速度惊人，但是在 GIS 应用方面还有一定局限，主要表现为地理数据的传输：目前，对于 GIS 数据的网络传输仍然有一些局限，由于 GIS 的数据通常容量较大，现在网络宽带的能力在中远距离的大量数据传输过程中，速度不够令人满意，这将会是网络技术在 GIS 发展过程中的一个瓶颈问题；GIS 网络软件的开发：网络技术给 GIS 技术的发展带来了更多潜力，但是到目前为止，GIS 软件工业界还没有充分地将这些潜力发挥出来，许多技术在 GIS 领域仍然处于研究和试验阶段，达到商业化、实用化还有一定的距离；网络技术在 GIS 中的有效使用：技术的发展只有在给人们带来利益时才有真正的价值。网络技术有巨大潜力，但是如何在 GIS 领域得到有效的使用，充分、恰当地发挥出它的潜能仍然是需要人们探索的问题。

2）GIS 标准化

今后 5-10 年是 GIS 界的主要标准化制定时期。GIS 发展到今天这样，能够在各种领域得到使用的盛况，人们不断意识到软件、硬件、数据等要素进行必要的标准化才能实现更有效、广泛地对 GIS 的使用。GIS 的标准化将在国际、国家、省、市、县和机构范围内多层次地进行，其内容可能包括到 GIS 的各个组成部分、各个操作过程、各种数据类型、软件硬件系统等。标准化的真正实现将使人们能在一个共同理解基础上共享信息和资源。

3）数据商业化

在西方社会，计算机硬件设备的生命周期通常为 3-5 年，计算机软件的生命周期一般为 7-15 年，而地理数据的生命周期则为几十年。地理数据的开发、更新和维护既费时又费力，在 GIS 界曾经有人统计过，GIS 硬件、软件和数据的造价比是 1：10：100，所以如何更有效地生产和维护地理数据将会是 GIS 未来面临的主要挑战之一。GIS 产生的主要目的之一是对于空间信息进行更好的管理和处理，GIS 空间分析功能实际上是使用现有的数据来产生新的数据，所以数据是整个 GIS 的操作对象。没有数据，则谈不上信息系统。如果数据问题能够解决，信息系统才有意义和价值，才能够真正运行。

4）系统专门化

目前，GIS 软件和系统还是被作为一个整体独立存在。许多软件提供全面的 GIS 功能可以在任何一种需要 GIS 的部门使用，没有具体专业领域的限制；而从使用 GIS 机构的角度上看，很多机构只是需要 GIS 软件中的部分功能，而目前 GIS 软件设置使得用户在购买 GIS 系统时往往要求整个软件一起购买。首先，从 GIS 用户方面考虑，GIS 可能不将作为一个独立的系统存在于机构内，而是作为机构整个管理和运作系统的一个部分，GIS 的各种功能将融合在与专业领域更直接的系统之中。其次，从 GIS 的工业发展角度上考虑，目前的 GIS 软件所提供的各种功能也将与各类专业软件系统融合起来，共同发展。软件的部件化是这个趋势的前兆，也为 GIS 软件的专业化做了必要的准备。将来的各类应用系统中，GIS 可能将作为一个必须的部分存在。

5）GIS 企业化

GIS 网络化的发展使得 GIS 在机构内部各部门之间更有效地进行通讯、交流和进行各种资源的共享。企业和机构可以从更高的层次上对 GIS 在企业中的使用进行统筹安排和计划，这种方式被称为“企业化 GIS”。企业化 GIS 对技术和管理人员有更高的要求，需要企业不断地对人员进行技术和管理培训。

6）GIS 全球化

网络技术的发展使得世界空间缩小，使人们之间的关系更加紧密；世界经济的发展也在要求人们建立一个更稳定、和谐的环境。在这个环境中 GIS 越来越成为一种有效的工具来帮助人们了解他们所生存和依赖的自然条件状况和社会变化状况。目前世界各国都在积极地发展和使用 GIS、制定有关地理信息的政策、开展国家的 GIS 项目，例如世界银行和其它的国际信贷组织都要求在它们资助的项目中使用 GIS 来辅助决策。GIS 的标准化对于它在国际范围内的推广和使用将起到促进作用，国际标准组织已经专门就地理空间技术从各个方面进行标准化制定和实施。

7）GIS 大众化

GIS 不仅在国际舞台上已经越来越受到人们重视，甚至在人们日常生活中也潜移默化地改变着人们的生活。以往人们需要使用地图来定向、定位和导航，而现在地图已经存储在数据库中；从一个地点到另一个地点的最佳路线轻而易举地就可以使用 GIS 系统得到；到一个新地方，不需要再费力寻找餐馆、旅店、娱乐中心、购物中心、银行、旅游景点等，GIS 就是最好的向导。

我们借助于外感官（我们意识的一种性质）表象给我们自己外面的对象，这些对象毫无例外的在空间里面。这些对象的形状、大小、以及它们相互间的关系是在空间里被规定的或能够在空间里被规定的。

空间不是一个从外部经验得来的经验概念。因为为使着某种感觉与我以外的某些东西发生关系，以及同样地为着我能把那些感觉表象为互相在外、互相靠近，从而不只是彼此不同，并且是在不同的地方，这样就一定要以空间观念为前提。

康德

第二章 从现实世界到比特世界

导读：本章从空间认知的角度讲述了对现实世界进行抽象的过程。空间认知属于行为地理学的范畴，它研究个体如何对现实世界进行认知，并在意识中编码的过程。在认识的基础上进行逐步抽象，最后得到数字化的空间数据，这正是 OpenGIS 九层抽象模型前五层所描述的内容。最后是空间数据库模型，进一步强调了该抽象过程。

本章的内容有助于加深对空间数据和空间信息的理解。

从现实世界到人的概念世界，再到数字世界，最后通过用户改造自然的活动反馈到现实世界，这中间三个阶段对应着地球信息科学的三个研究领域，即地理认知模型的研究、地理概念计算方法的研究以及地理信息与社会的研究。如图 2-1 所示。

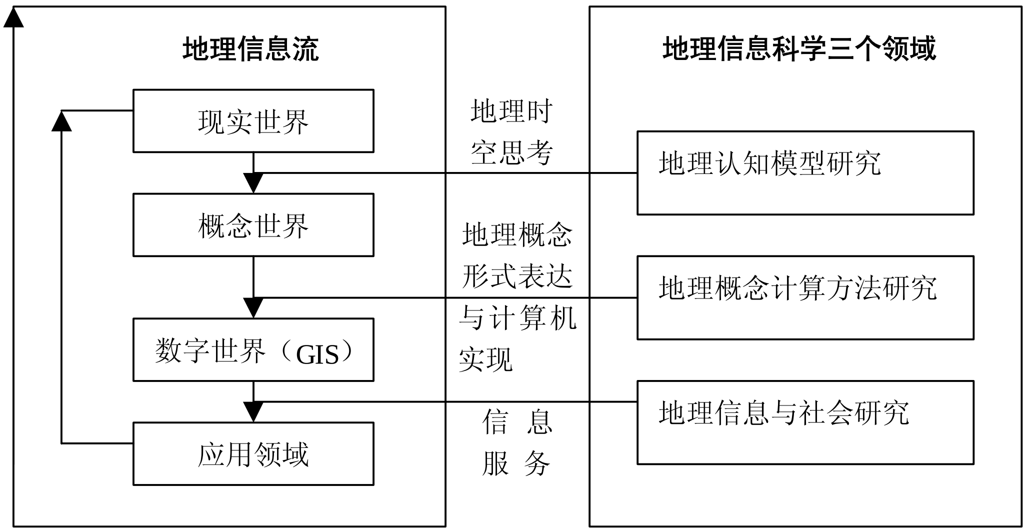


图 2-1： 地理信息流与地理信息科学三个领域

1．对现实世界的地理认知

地理环境是复杂多样的，要正确地认识、掌握与应用这种广泛而复杂的信息，需要进行去粗取精、去伪存真的加工，这就要求对地理环境进行科学的认识。对于复杂对象的认识是一个从感性认识到理性认识的一个抽象过程。对于同一客观世界，不同社会部门或学科领域的人群，往往在所关心的问题、研究的对象等方面存在着差异，这就会产生不同的环境映象。

1．1 认知的含义

认知属于心理学的范畴。认知理论早在 70 年代就被引入地图学，并用“刺激——反应”的关系模式来研究用图者在读图时的心理——物理反应。始于 90 年代的地图制图可视化研究也应用了认知概念，使

“刺激——反应”模式的研究向纵深推进了一步，地图制图可视化的实质是探讨地图信息传输过程中人对空间客体的认识。根据空间信息分析以及空间信息可视化的需要，认知应该是知觉、注意、表象、记忆、学习、思维、语言、概念形成、问题求解、情绪、个性差异等有机联系的信息处理过程。

在简单的认知系统中，感知系统从外界环境输入刺激、进行变换整合，提取刺激特征加以组合，输出经编码的物理刺激。长时记忆系统将输入的经编码的物理刺激与系统只能感知的信息进行比较与模式匹配，一部分信息被激活，这部分被激活的信息送入短时记忆系统。短时记忆系统对输入的有限信息进行精细加工，通过反映系统来输出动作、语言、表情等。中枢处理器进行系统控制和处理，决定目标先后次序和监督当前目标执行，图 2-2 描述了空间认知的框架。

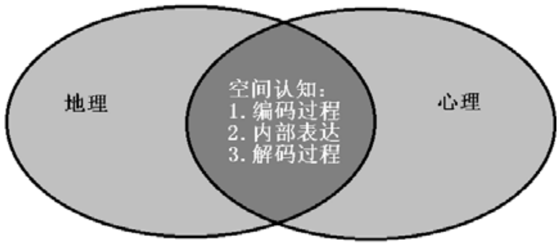


图 2-2：空间认知

空间认知

地理学家和心理学家对空间进行了大量研究，得到了许多有趣的结论，它们成为行为地理学中研究人的空间决策的基础; 在 GIS 中研究空间认知，可以指导建立更加符合人类思维的空间模型和信息表现方式。下面介绍的是与空间认知有关的知识。

1）个体的空间认识的形成和发展

J.Piaget 认为，空间认识具有四个层次：感觉运动（Sensorimotor）的层次，它以行动的表象为依据；前运算期（Pre-operational）的层次，其基础是从记忆中获得的有关客观世界的知觉的映象；形象运算（Concrete Operational）的层次，其中允许符号化、系统化的环境心理表象；形式运算（Formal Operational）的层次，在该层次中，个体具有假设——演绎推理能力，能够处理独立于行动、客体、空间的抽象的空间概念。图 2-3 表示了个体成长过程中的空间认识的发展。

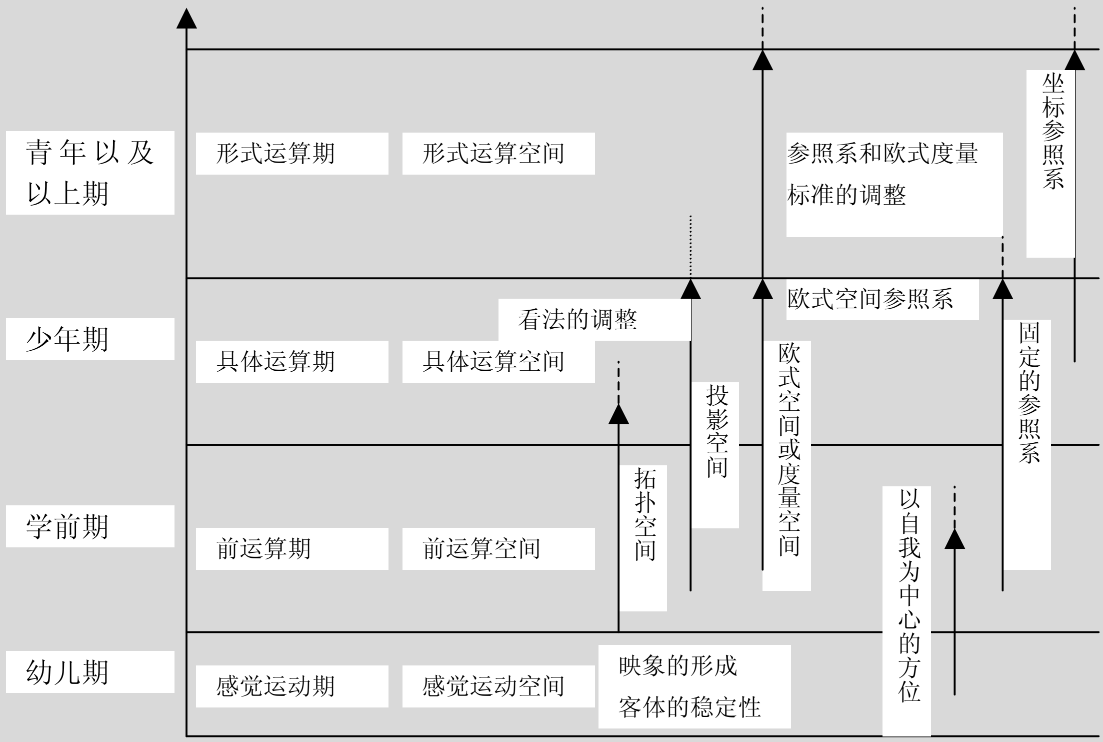


图 2-3：空间意识的个体发育

2）意识中的空间信息编码[Elisabeth S. Nelson]

关于个体意识中的空间信息的编码方式，主要有三种理论：命题理论（Propositional Theory），图像理论（Imagery Theory）和双重编码理论（Dual-coding Theory）。

（2．1）命题理论

一个命题，是表达了对象之间关系的概念结构。命题必须符合以下规则：它必须是抽象的；它必须与一个真值关联；它必须符合一定的形式化规则。按照命题理论，空间编码形如以下的单句“俄克拉荷马在德克萨斯以北”，命题理论的关键是认为其命题只是对抽象的意义编码，而非信息的自然特征。

（2．2）图像理论

图像理论认为，在空间认知中，图像是其关键构成。其理论的基础是一个图片的隐喻，认为个体感知并处理空间信息，然后将其组织成为简单的、更加有序的形式。这些信息被记忆，并且可以在需要时重新编排生成图像。

（2．3）双重编码理论

双重编码理论认为语言的和视觉的信息被独立的处理，但是相关的处理系统相互连接。一个图像系统组织各种简单的图像，形成层次结构并且可以输出为空间的形式；而语言系统并行地处理各种非视觉信息，并组织成为更高层次的、连续的结构。

地理信息系统就是信息加工系统，即输入信息，进行编码，存储记忆、做出决策、输出结果。这就是环境信息流在人的大脑中的处理过程被地理信息系统所模拟和复制的原因，当然，认知的目的在于求解问题，即找到问题的解决方案。认知操作包括常规问题求解和创造性问题的求解。

1．2 环境映象与模型

环境映象可以看作是被学习到的稳定的思维概念，它总结了个人对环境的认识、评价和选择能力，对任何人与环境相互作用的调查研究来说，理解大脑处理使用什么样的信息就显得至关重要。人与环境相互作用的关系见图 2-4 所示。

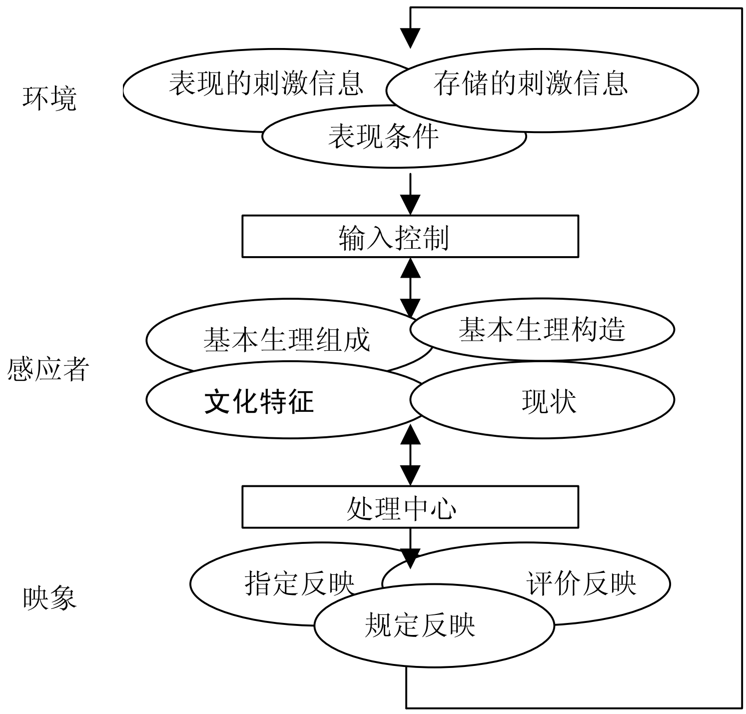


图 2-4：人与环境的相互作用关系图

在环境映象研究中，如何从个人那里提取对大尺度环境有意义的信息？如何展现该信息，以便使个人的认知范围和程度明朗化？如何分析这些信息？最为普遍的方法是绘制描述环境要素的简图，通过描述环境要素的序列、要素之间的联系和差异、映象的类型、细节和规模的变化等来提供信息，即构造模型。在科学研究中，经常用模型作为实际对象的代替物或模拟物，它们与被研究的对象之间存在某种程度的一致或相似的映射关系：借助模型所获得的结论和推论又可转用于实际对象上。地理学中所采用的表现模型多

种多样，例如各种地图、航空和航天遥感图像、统计图表、剖面图等，但是地图是地学中使用最普遍、也是最重要的模型。从模型的特性和实质来看，地图是客观世界的形象/符号/概括模型。单幅地图是模拟客观实际的某一方面的模型，而系列地图和地图集属于高层次的模型，即地理系统的模型，它们模拟的是自然要素和社会经济要素的总体，反映出这些要素的相互联系、相互作用、解析特征与综合效果、现状与未来发展趋势，以及它们所形成的复杂巨系统的功能特征等。

1．3 地图认知模型

地图认知模型分为地图编制与设计者的认知模型和地图使用者的认知模型（图 2-5 和图 2-6）。

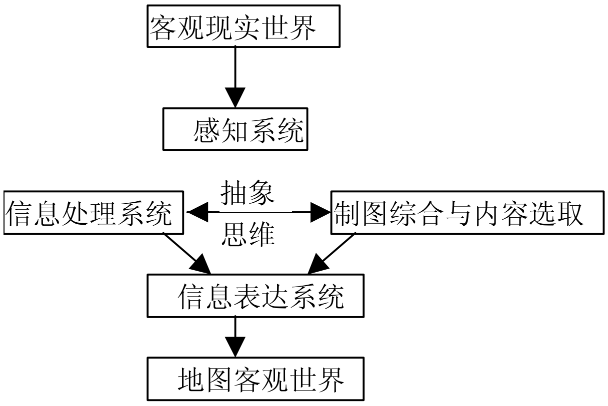


图 2-5：地图编制者的认识模型

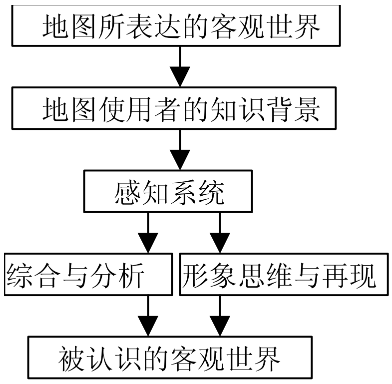


图 2-6：地图使用者的认识模型

制图者的认知模型强调对所表达事物和现象的认知，以及对表达内容的表现形式的认知。制图者认知的目的是指导制图者选取最主要的制图内容与最适合的表现形式，以高效地传输空间信息（图 2-7）。

地图使用者的认知是在已有地图的基础上，结合读图者的自身的空间知识与背景，完成对地图对象的认知，从而间接达到认知客观世界的目的。

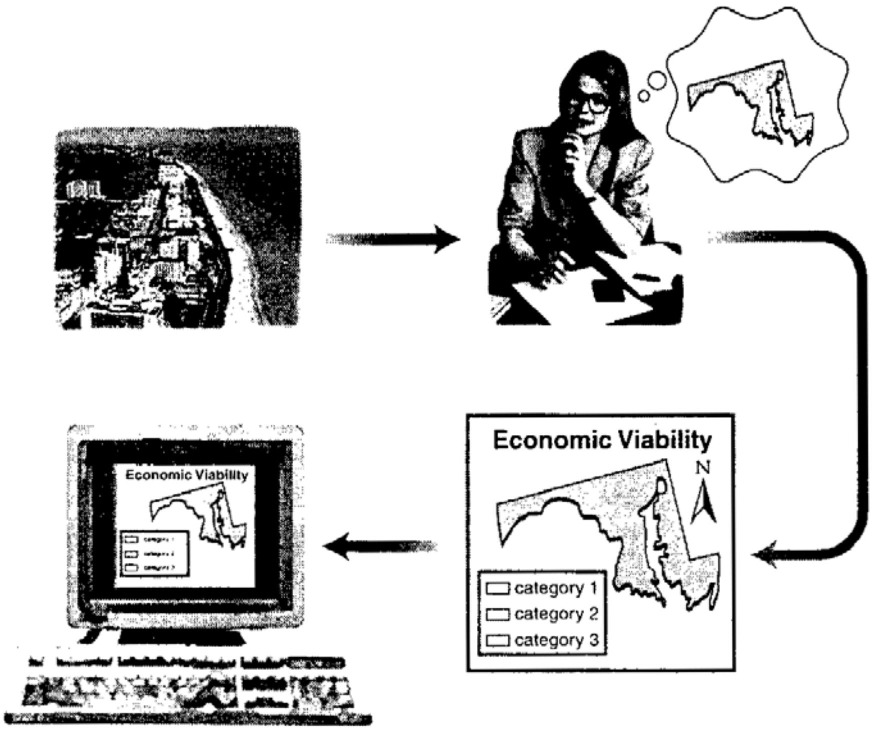


图 2-7：制图活动过程

1．4 地理客体的科学认知

地理认知是地理信息传输过程中一个子系统。它偏重于心理感知和分析，认知者既感知图上明显的信息也挖掘潜在的信息，不仅仅是探测、识别或区分信息，更要主动地解译信息，形成对客观世界的整体认识。从地图学者在编制地图时的地理认知，到用图者在读图时的地理认知，这整个过程反映了人对地理课题的认识由浅入深的特点。因为从原始制图资料到地图再到新地图的地理信息传输过程，正是人们对地理事物的认知深度的螺旋式上升过程。地理认知不仅是地理模型的基础，而且也是制图概括的基础。

在地理认知的基础上对真实世界的信息进行抽象和概括，形成地理信息系统中的模型，在一定比例尺下显示地理要素的分类、分级和空间图形格局。地图模型本身就是一个主观与客观相统一的过程，GIS 专家在地理认知中不仅全面、综合地分析和理解地理环境，而且其认知中还包括了对一些客观的模拟标准和规范的接受或认同。

此外，地理认知贯穿于制图概括的整个过程。具体来看，它表现在地图设计阶段，对制图概括原则、内容、分类分级指标的“构思”上，以地理认知中所获得的关于地理环境的系统功能、层次结构、各个要素的组合关系等的知识为指导。在地图编绘阶段，对地图要素的选取、数量和质量概括、图形简化等“构图”操作上，在处理符合概括指标的要素或图形时能够进一步体会地理认知的控制和指导作用，做到心中有数，操作正确，不至于误解概括指标和规则。

1．5 地图是客观世界的形象——符号——概括模型

1．5．1 地图是客观世界的形象模型

地图与其它地学模型的主要区别和优点在于它具有形象性，能对实际对象做出完整的、清晰的和直观的图形描述和说明。“地图是地表空间关系和空间形式的视觉图解表象”。

1．5．2 地图是客观世界的符号模型

地图区别于其它许多图示模型的又一特征是它采用专门设计和事先规定的符号来反映地物、现象和地理过程，并表示它们的位置、质量特征和数量特征。各种符号的组合便构成一种地图形象，而多种地图形象的汇集便组成一幅完整的地图图形。地图符号的作用并不仅仅局限于传输信息，它们还是记录知识、使知识定型并系统化的强大工具。

1．5．3 地图是客观世界的概括模型

地图不是客观对象的完全再现，而是通过地图制图工作者脑和手的“过滤”、概括和抽象出来的模型。对地物进行取舍、图形化简、数量和质量概括、地物协调，以及夸大表示等处理，是地图学者对地物进行深刻理解、综合分析并抽象概括等一系列复杂的科学思维和创造性劳动的结果。因此，对制图对象进行地图概括是地图作为客观实际模型的一个最重要的特征。

2．现实世界的抽象

对地理对象的抽象过程通常认为有 9 个层次[OGC]，在这九个层次之间通过 8 个接口与它们连接，定义了从现实世界到地理要素集合世界的转换模型。这 9 个层次依次为现实世界（Real World）、概念世界（Conceptual World）、地理空间世界（Geospatial World）、尺度世界（Dimensional World）、项目世界（Project World）、点世界（Points World）、几何体世界（Geometry World）、地理要素世界（Feature World）以及要素集合世界（Feature Collection World）（图 2-8）。连接它们的 8 个接口分别为认识（Epistemic）接口、GIS 学科（GIS Discipline）接口、局部测度（Local Metric）接口、信息团体（Community）接口、空间参照系（Spatial Reference）接口、几何体结构接口、要素结构接口及项目结构接口。其中前五个模型是对现实世界的抽象，并不在计算机软件中被实现；后四个模型是关于真实世界的数学的和符号化的模型，将在软件中被实现*。

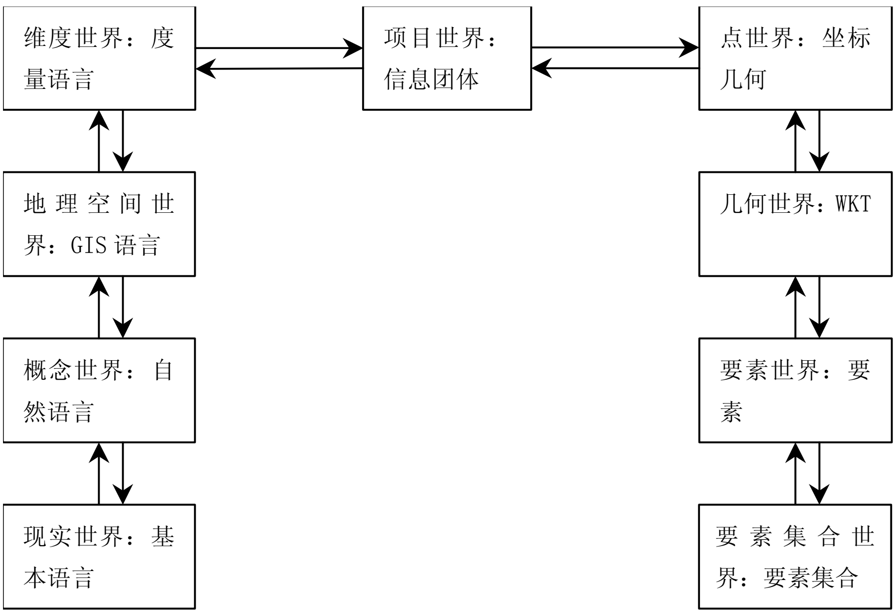


图 2-8：OpenGIS 的九层模型

* 因为九层模型后面四个层次与软件实现有关，涉及具体的数据结构，本章不做介绍，可以参照“空间数据模型”和“空间数据管理”两章的内容。本章只是介绍前五个模型——对现实世界进行抽象，得到数字化的、GIS可以管理的模型。

2．1 现实世界

现实世界是所有事物（Fact）的集合，无论人们是否知道这些事物。根据事物的本质，人们可以认识理解现实世界中的事物。图 2-9 表示了人类生活于其中的现实世界。象云一样的纹理结构占据了图形的绝大部分，代表了人们所不了解的事物，而且它们造成了宇宙的混沌状态。人们只知道一些所熟悉的事实，其中的一些被绘制在图中。

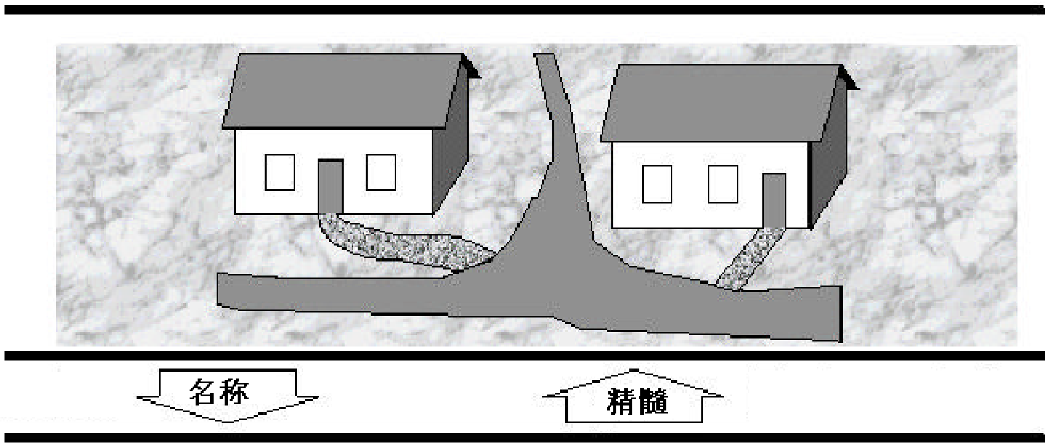


图 2-9：现实世界

2．2 概念世界

概念世界是人类自然语言的世界，人类了解且认识其所命名的事物，因此这些事物构成了“语言的世界”。在图 2-10 的概念世界中，表示宇宙混沌状态的云并不存在，因为这些在自然语言内容中通常是不可见的。示意图显示了容易识别的事物：门、路、砖、屋顶、房屋等。用这种方法，可以返回到真实的世界中，抽象出一个事实的本质，称之为精髓(Pith)。由于可以给出所知道的事物的名字，而且能够感受到这些相同的已知事物的本质，所以称真实世界与概念世界之间的交互为认识接口（Epistemic Interface）。

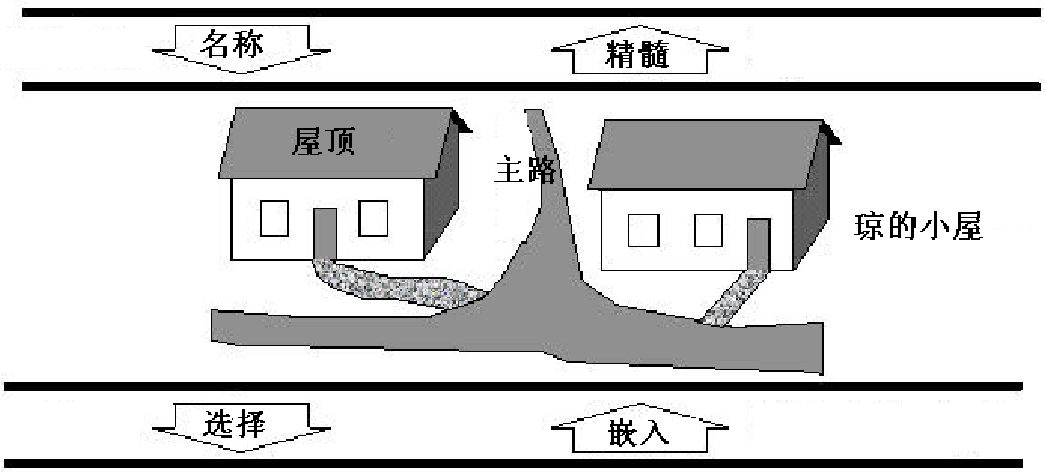


图 2-10：概念世界

对于 GIS 来讲，自然语言的概念世界并不是充分抽象的，在 GIS 中只有概念世界中一个简化的子集才是兴趣所在。这个子集叫做地理空间世界，人们与概念世界的交互的方法叫选择。

图 2-11 中，有三种实体类型。每个由一个矩形表示，并且每个用其上部的名字所表示。矩形之间的线代表着实体之间的联系，每条线在末端带有一个作用名以解释这种联系。菱形代表聚合（Aggregation），

例如，概念世界的存在依赖于真实世界的存在。实心圆意味着在联系的那个末端存在着一个集合（并不是一个单个的对象），例如，每个概念世界嵌入（Embed）一系列不同的地理空间世界。

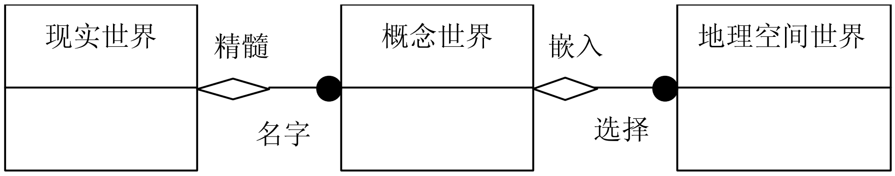


图 2-11：现实世界与地理空间世界的联系

2．3 地理空间世界（Geospatial World）

从事 GIS 的技术人员都习惯于把世界看作一个抽象的、几乎是具有卡通特性的世界。这是由于在概念层次的世界充满了复杂的形状、样式、细节。这些复杂性在地理空间世界中被消除，并用简单的、浅显的抽象来代替，这些抽象通常在时间以及空间上都是静态的。通过地理空间世界的抽象，河流是被看作线，地形被看作等高线多边形的简化，而森林被看作多边形。

前面描述的概念世界在图 2-12 中以卡通的方式在地理空间世界水平上被重画。图 2-12 是以透视的方式绘制的，但是地理空间世界通常是从一个“顶点”来观察，即垂直地从上向下看。注意在图中，一些特征已经消失，另外一些已经变得大大简化。例如，一些窗户、墙、建筑物的房顶已经消失。这是由于它们并不是 GIS 世界视点的兴趣所在。它们已经成为 GIS 意识中所不可见的。当然，并不存在一个通用的定义——准确地讲，什么样的特征对于一个 GIS 技术人员来讲是其兴趣所在，也许有时一个屋顶可能会引起兴趣。卡通仅仅表明地理空间世界是概念世界中的一个子集和一个简化，在地理空间世界中所讲的语言是 GIS 学科语言。在图中，房屋的地基还保持着，尽管在概念世界中，部分地基被隐藏在其它要素后面。从 GIS 的观点来看，一个建筑物的所有的地基都是可见的——尽管有些是看不见的。每个 GIS 实现都有特定的规则，这些规则规定了在地理空间世界中什么样的特征被识别以及它们是如何从概念世界中被简化的。例如，一个规则可以把一个砖房简化为一个三维的多面体；然而具有另一种表面物质的一个房屋被简化为它的地基多边形。简单地讲，在概念世界中不可见的事物在地理空间世界中就变得可见了，因为这些事物在 GIS 中是特殊兴趣所在。概念世界和地理空间世界之间的交互称为 GIS 学科接口；从概念空间进行的交互方法叫做选择。为了转换从地理空间到概念世界的这种交互，人们可以采用嵌入（Embed）方法，这种方法把 GIS 感兴趣的内容放置在概念世界适当的语境（Context）中。

在地理空间世界中所认识的要素通常有一个自然的维度：0、1、2 或者 3，这取决于它们是否被看成点、线、面、体。此外，根据二元拓扑关系（如包含、相邻或分开），它们还具有另外的量度。下一个层次的抽象识别了要素固有的维度和尺度（Dimensionality and Metrics）特性，因此叫做维度世界，可以通过在欧氏空间中进行工具测量获取要素的尺度。

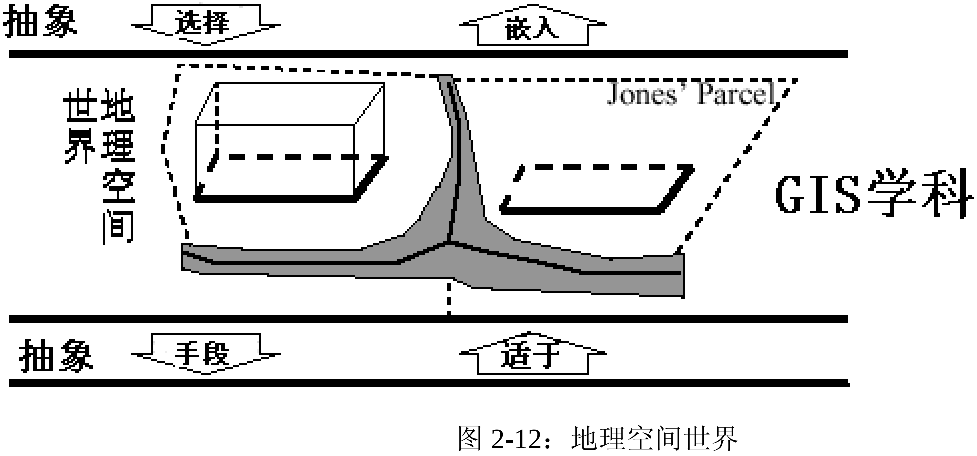


图 2-12：地理空间世界

2．4 维度世界（Dimensionality World）

维度世界是地理空间世界的一个抽象，其中包括一些测量工具，如卷尺测量和指南针，在这样的水平上所认识到的事实包括一元（Unary）关系（如一个弧段的长度）和二元（Binary）关系（如两点之间的距离），这些关系本身就是各种要素的抽象。

维度世界与地理空间世界的接口叫做适合（Fit）。两个电话线路杆之间的距离是属于维度世界。这个长度线适合于地理空间世界中所见到的长度跨度，在图 2-13 中包括了在维度空间所表现的一些抽象。

维度世界是现实世界抽象的最后一个。下一个抽象叫做一个项目世界，该世界只发生在一个具体的实现中，每一个实现都是针对一个特殊的 GIS 学科或分学科。在每一个实际的实现中，只有维度世界中的一个子集得到识别。通常这个子集是由研究区域的范围和被测量的特定的现象所决定的。

在项目世界的层次中，空间参照系统（Spatial Reference System）的概念被引入，最常见的参照系统是环绕地球表面建立的坐标系统（经纬度）；此外，还有其它的间接的参照系统，例如线性参照系统可以用一个参数标识出一条线（如高速公路）上的一个点位。无论采用哪种坐标参照系统，都可以确定维度世界中的要素的每个“角点”的坐标。维度世界和项目世界之间的接口称为信息团体接口，从维度世界调用该接口的方法叫做编码（Codify），其结果是每个“角点”的坐标以一组数值表示；相反的，从项目世界调用的方法称为定位，它确定了每个要素与其他要素的相对关系。

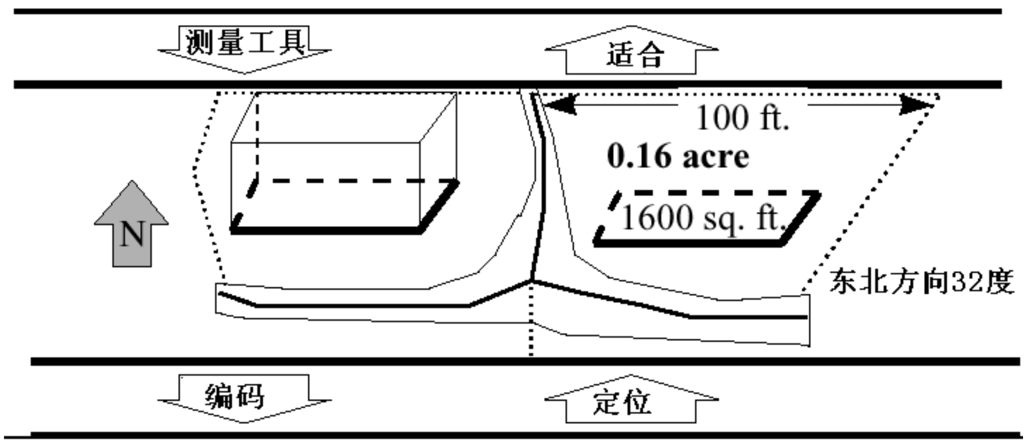


图 2-13：维度世界

2．5 项目世界（Project World）

有两种常用的方法来对地理空间要素建模。第一个模型定义了点、线、多边形的一个要素的空间范围，以及来自于所熟知的一系列类型的几何基本单元（Primitive），在这种方式下的要素叫做“几何体要素（Features with Geometry）”。

第二种称为覆盖（Coverage），影像是该模型的一个特殊的例子。几何体要素和覆盖是紧密相关的，但是在概念上截然不同*。

要素模型用于对现实世界中的地物对象建模，如道路、城市等等；而覆盖模型则对现象建模，包括：温度、土壤分布等等。一个要素具有多种属性，如空间位置属性，空间关系属性，描述属性，时间属性等等，覆盖也可以视为要素的一种属性。

GIS 并不只是一个学科，它还是地理空间信息表现的一门语言，其中这些信息源于许多与地理学科相关的学科，如森林管理、土壤制图、运输模型、地籍管理等等，这些学科中的每一门学科都有许多子学科，一个 GIS 项目可以包含任何这些学科的组合。

---

* “空间数据模型”一章中描述的“要素模型”和“场模型”对应于这两种技术。

正是项目世界语言的多样性导致了 GIS 信息存储之间相交互的最复杂的问题。这是人为造成地理空间世界分割的原因。但是，如果对语言结构进行充分地统一化，这种情况是可以管理的。

地理信息团体是指共享数据的用户群，他们属于不同的专业领域，可以是数据使用者，也可以是数据提供者。地理信息团体将整个地理空间世界中的一个特殊的子集看成是一个抽象。在三个不同的应用中，上述的例子被抽象为三个不同的项目世界模型。它们显示在图 2-14 中，分别反映了一个制图员、一个地籍管理人员和一个道路管理人员视角的项目世界。

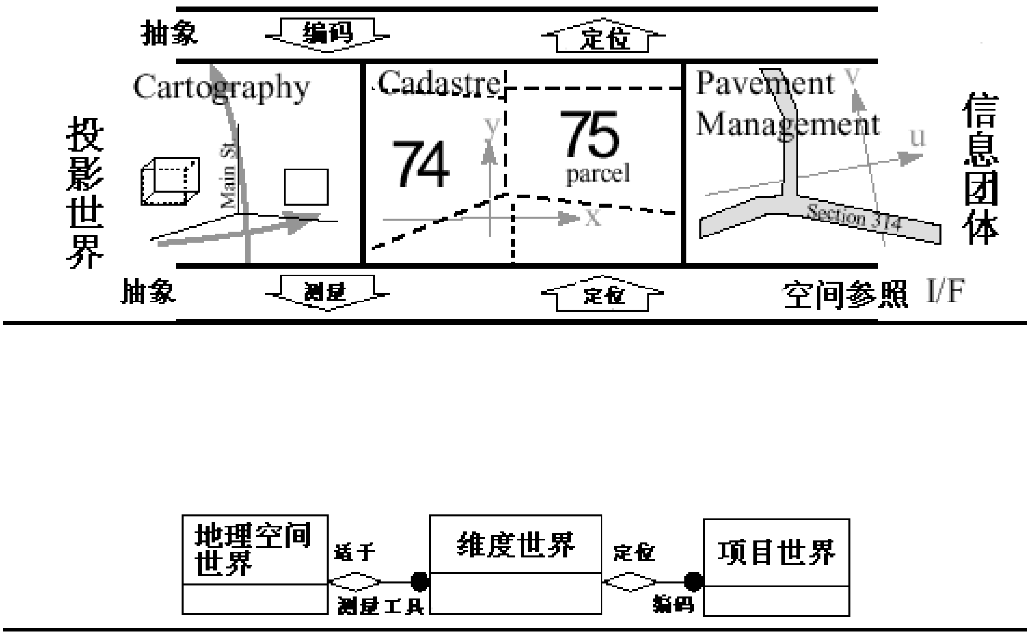


图 2-14：项目世界

3．比特世界

3．1 比特世界

地理信息系统以数字世界表示自然世界，现实世界与数学模型之间的关系见图 2-15。

在计算机中，现实世界是以各种符号形式来表达和记录的，计算机在对数字和符号这些符号进行操作时，又将它们表示为二进制形式（比特世界）。因此，基于计算机的地理信息系统不能直接作用于现实世界，必须经过对现实世界的数据描述这一步骤。模型是对现实世界的简化表达，是将系统的各个要素通过适当的筛选，用一定的表现规则描写出来的简明的映象。

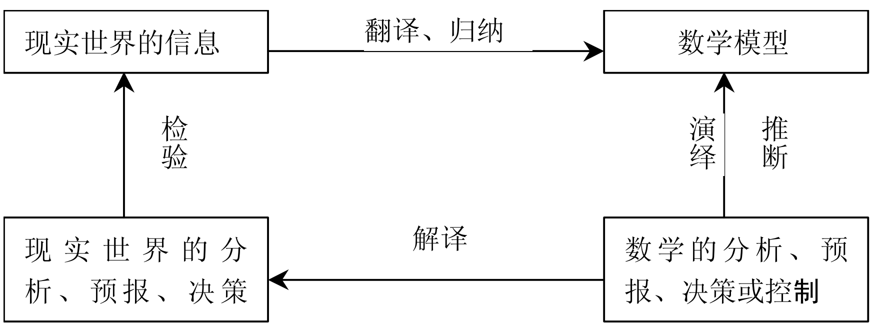


图 2-15：现实世界与数学模型的关系

一幅地图是一个符号模型，因为它是通过制图学家处理后得到现实世界的简化描述；存储数字地图的

计算机文件也是一种符号模型，它以数字代码来表现图形符号。一幅数字地图的产生不仅需要选择所要表现的物体，还要进一步考虑如何对表达它们的数据进行组织。如果数据的组织规则没有很好地建立起来，则一幅数字地图除了对生产这些数据的个人或组织有用以外，对于其他人是没有用的。

数据是对现实世界状况的数字符号记录，信息是经过重新组织的，能揭示现实世界内在机理的并有利于研究工作的数据。如果数据不以空间属性表来组织，则很难从空间数据中抽取出空间信息。由于计算机的数字化特征，数据项必须是离散的以便于进行数字处理和操作，因此，地理空间也必须离散化的表达。数据建模是指把现实世界的数据组织为有用且能反映真实信息的数据集的过程。根据一定的方案建立的数据逻辑组织方式叫数据模型。数据建模过程分为三个步骤：首先，选择一种数据模型来对现实世界的数据进行组织；然后，选择一种数据结构来表达该数据模型；最后，选择一种适合于记录该数据结构的文件格式。可见，一种空间数据建模可能有几种可选的数据模型，而每一种数据结构又可能有多种文件格式进行存储。空间数据可依据它们的采集方式、存储方法、使用目标等，用不同的数据模型进行组织。如GIS中最常用的数据组织方式为矢量模型和栅格模型。在矢量模型中，用点、线、面表达世界，在栅格模型中用空间单元或像元来表达世界。

3．2 模型的作用

在地理信息系统中，模型，尤其是数学模型起着十分重要的作用。由于模型是对客观世界中解决各种实际问题所依据的规律或过程的抽象或模拟，因此能有效地帮助人们从各种因素之间找出其因果关系或者联系，有利于问题的解决。模型的建立是数学或技术性的问题，但它必须以广泛、深入的专业研究为基础，专业研究的深入程度决定了所建模型的质量与效果，而模型的质量和数量又决定了系统中数据使用的效率和深度。大量模型的发展和应用，实际上集中和验证了该应用领域中许多专家的经验和知识，这无疑成为一般地理信息系统向专家系统发展的基础。

3．3 GIS 空间数据建模

GIS 是专门用于采集、存储、管理、分析和表达空间数据的信息系统，它既是表示、模拟现实空间世界和进行空间数据处理分析的工具，也是一门关于空间信息处理分析的科学技术。就 GIS 的工具特性而言，它是为人们采用数字形式表示和分析现实空间世界提供了一系列空间操作和分析的功能，包括综合地存储管理人们研究和解决空间问题所需的各种空间数据；根据用户的要求查询有关的空间分布信息，进行各种统计量算、列表制图；根据规划、管理、生产的需要，进行多因素的综合研究、决策方案的模拟优化等。所以，GIS 一方面应为用户采用数字形式表示和分析空间现象或问题提供空间数据建模及分析操作的手段，另一方面应向用户提供友好的使用界面，便于用户进行空间数据建模、查询、分析等。

空间数据建模的基本任务是，针对所研究的空间现象或问题，描述 GIS 的空间数据组织，设计 GIS 空间数据库模式，这包括定义空间实体及其相互间关系，确定数据实体或目标及其关系，设计在计算机中的物理组织、存储路径和数据库结构等。这项工作是以空间数据模型的理论为指导的。空间数据模型是关于现实世界中空间实体及其相互间联系的概念，为描述空间数据组织和设计空间数据库模式提供了基本的方法。

一般而言，GIS 空间数据模型由概念数据模型、逻辑数据模型和物理数据模型三个有机联系的层次所组成。其中概念数据模型是关于实体及实体间联系的抽象概念集，逻辑数据模型是表达概念数据模型中数据实体（或记录）及其间关系，而物理数据模型则是描述数据在计算机中的物理组织、存储路径和数据库结构，三者之间的相互关系如图 2-16 所示。

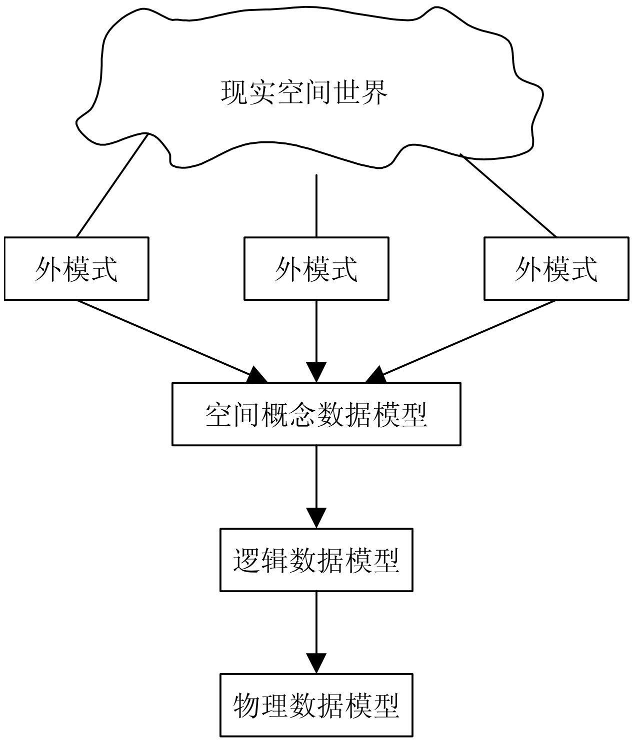


图 2-16：空间数据模型的三个层次

3．4 GIS 空间数据模型的概念与分类

3．4．1GIS 空间概念数据模型

由于职业、专业等的不同，人们所关心的问题、研究对象、期望的结果等方面存在着差异，因而对现实世界的描述和抽象也是不同的，形成了不同的用户视图，称之为外模式。GIS 空间数据模型的概念模型是考虑用户需求的共性，用统一的语言描述和综合、集成各用户视图。目前广为采用的数据模型是基于平面图的矢量数据模型和基于连续铺盖的栅格数据模型。

3．4．2 空间逻辑数据模型

逻辑数据模型是根据前述的概念数据模型确定的空间数据库信息内容（空间实体及相互关系），具体地表达数据项、记录等之间的关系，因而可以有若干不同的实现方法。一般来说，可将空间逻辑数据模型分为采用结构化模型和面向操作的模型两大类。

1）结构化逻辑数据模型

结构化模型是显式表达数据实体之间关系的树形结构。其中的层次数据模型是按树型结构组织数据记录，以反映数据之间的隶属或层次关系。网络数据模型是层次数据模型的一种广义形式，是若干层次结构的并，其优点是能反映现实世界中极为常见的多对多的联系，缺点是复杂。一般而言，结构化模型能直接地反映现实世界中空间实体之间的联系。

2）面向操作的逻辑数据模型

关系数据模型是用二维表格表达数据实体之间的关系，用关系操作提取或查询数据实体之间的关系，因此称之为面向操作的逻辑数据模型。其优点是灵活简单，但表示复杂关系时比其它数据模型困难；当数据构成多层联系时，存储空间利用效率较低。当前的一种发展趋势是将两者的优点集中起来，形成新的或改进的逻辑数据模型，如扩展的网络模型。

3．4．3 物理数据模型

逻辑数据模型并不涉及最底层的物理实现细节，但计算机处理的是二进制数据，必须将逻辑数据模型转换为物理数据模型，即要设计空间数据的物理组织、空间存取方法、数据库总体存储结构等。

1）物理表示与组织

层次逻辑数据模型的物理表示方法主要有物理邻接法、表结构法、目录法。网络数据模型的物理表示方法主要有变长指针表、位图法、目录法等。关系数据模型的物理表示是用关系表进行的。物理组织主要是考虑如何在外存储器上以最优的形式存放数据，通常要考虑操作效率、响应时间、空间利用和总的开销。

2）空间数据存取

数据库的“存”是指从内存写一块到外存，“取”指从外存写一段到内存。常用的存取方法有：

第一种，文件结构法：包括顺序结构（如二分查找，插值查找）、表结构（线性表，倒排表）和随机结构。

第二种，索引文件：它是提高数据存取效率的基本方法。对索引的插入、删除等只涉及到索引记录本身，而对数据记录的操作要看具体的数据组织策略。如果索引本身很大的话，就要对索引文件再索引，建立多级索引，如 B 树、B+树等。B 树是基于主关键字的索引，若要根据次关键字进行索引，必须建立倒排索引表。但是，如果这种基于次关键字的搜索是主要操作的话，这类索引就不适合了。

第三种，点索引结构：由于 B 树在进行基于次关键字的搜索时是不适合的，为此，将空间定位数据及其属性看作是多维空间中的点，采用栅格索引、KD 树、四叉树、R 树等多维点索引结构进行索引。目前空间存取方法及查询优化仍是 GIS 研究中的一个重要的课题。

藿乐威尔田园的真正迷人之处，在我看是：它的遁隐之深，离开村子有两英里，离开最近的邻居有半英里，并且有一大片地把它和公路隔开了；它傍着河流，据它的主人说，由于这条河，而升起了雾，春天就不会下霜了。

梭罗

在地球表面的任何地方都存在着垂直的和水平的两种关系：垂直关系把同一个地方的不同要素联结起来，而水平关系则把不同地方的各种因素联结起来。这两种关系的相对重要性随时代的变化而有所不同…正是这双重的关注，甚而至于这两种关系的结合，才为地理学提供了独特性和完整性。

R. J. 约翰斯顿

第三章 空间数据模型

导读：本章描述的是整个 GIS 理论中最为核心的内容。为了能够利用信息系统工具来描述现实世界，并解决其中的问题，必须对现实世界进行建模。对于地理信息系统而言，其结果就是空间数据模型。空间数据模型可以分为三种：

场模型：用于描述空间中连续分布的现象；

要素模型：用于描述各种空间地物；

网络模型：可以模拟现实世界中的各种网络；

在各种模型中，又介绍了相关的概念，如空间划分，空间关系，以及拓扑关系的形式化描述——9 交模型等。

最后讲述了普通的二维数据模型在空间上和时间上的扩展，时间数据模型和三维数据模型。

值得注意的是，本章谈到的场模型和要素模型类同于后面提及的栅格数据和矢量数据，但是前者是概念模型；后者是指其在信息系统中的实现。

1. 空间数据模型的基本问题

人类生活和生产所在的现实世界是由事物或实体组成的，有着错综复杂的组成结构。从系统的角度来看，空间事物或实体的运动状态（在特定时空中的性状和态势）和运动方式（运动状态随时空变化而改变的式样和规律）不断发生变化，系统的诸多组成要素（实体）之间又存在着相互作用、相互制约的依存关系，表现为人口、物质、能量、信息、价值的流动和作用，反映出不同的空间现象和问题。为了控制和调节空间系统的物质流、能量流和人流等，使之转移到期望的状态和方式，实现动态平衡和持续发展，人们开始考虑对其中诸组成要素的空间状态、相互依存关系、变化过程、相互作用规律、反馈原理、调制机理等进行数字模拟和动态分析，这在客观上为地理信息系统提供了良好的应用环境和重要发展动力。

1．1 概念

地理数据也可以称为空间数据（Spatial Data）。地理空间是指物质、能量、信息的存在形式在形态、结构过程、功能关系上的分布方式和格局及其在时间上的延续。地理信息系统中的地理空间分为绝对空间和相对空间两种形式。绝对空间是具有属性描述的空间位置的集合，它由一系列不同位置的空间坐标值组成；相对空间是具有空间属性特征的实体的集合，

由不同实体之间的空间关系构成。在地理信息系统应用中，空间概念贯穿于整个工作对象、工作过程、工作结果等各个部分。空间数据就是以不同的方式和来源获得的数据，如地图、各种专题图、图像、统计数据等，这些数据都具有能够确定空间位置的特点。

空间数据模型是关于现实世界中空间实体及其相互间联系的概念，它为描述空间数据的组织和设计空间数据库模式提供着基本方法。因此，对空间数据模型的认识和研究在设计GIS 空间数据库和发展新一代 GIS 系统的过程中起着举足轻重的作用（图 3-1）。

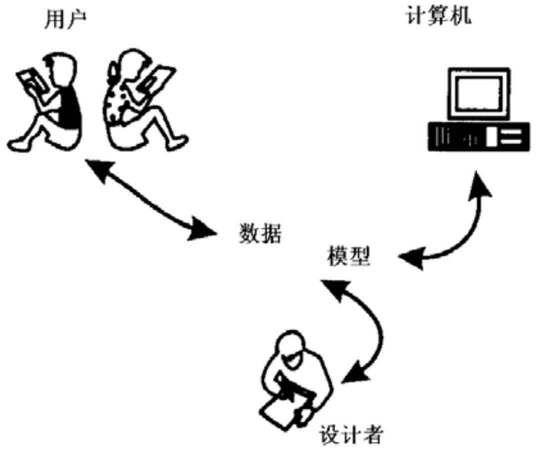


图 3-1：概念数据模型

1．2 空间数据模型的类型

在 GIS 中与空间信息有关的信息模型有三个，即基于对象（要素）（Feature）的模型、网络（Network）模型以及场（Field）模型。基于对象（要素）的模型强调了离散对象，根据它们的边界线以及组成它们或者与它们相关的其它对象，可以详细地描述离散对象。网络模型表示了特殊对象之间的交互，如水或者交通流。场模型表示了在二维或者三维空间中被看作是连续变化的数据。

有很多类型的数据，有时被看作场，有时被看作对象。选择某一种模型而不选择另外一种模型主要是顾及数据的测量方式。如果数据来源于卫星影像，其中某一现象的一个值主要是为区域内每一个位置提供的，如作物类型或者森林类型可以采用一个基于场的观点；如果数据是以测量区域边界线的方式而且区域内部被看成是一致的，就可以采用一个基于要素的观点；如果是将分类空间分成粗略的子类，一个基于场的模型可以被转换成一个基于要素的模型，因为后者更适合于离散面的或者线的特征的度量和分析。

1．3 GIS 空间数据模型的学术前沿

时空数据模型、三维数据模型、动态空间数据结构、分布式空间数据管理、空间存取方法、GIS 设计的 CASE 工具等是目前国际上 GIS 空间数据模型研究的学术前沿。

1．3．1 时空数据模型

时空数据模型的核心问题是研究如何有效地表达、记录和管理现实世界的实体及其相互关系随时间不断发生的变化。这种时空变化表现为三种可能的形式，一是属性变化，其空间

坐标或位置不变；二是空间坐标或位置变化，而属性不变，这里空间的坐标或位置变化既可以是单一实体的位置、方向、尺寸、形状等发生变化，也可以是两个或两个以上的空间实体之间的关系发生变化；三是空间实体或现象的坐标和属性都发生变化。当前时态 GIS 研究的主要问题有：表达时空变化的数据模型、时空数据组织与存取方法、时空数据库的版本问题、时空数据库的质量控制、时空数据的可视化问题等。

1．3．2 三维空间数据模型

国际上关于三维空间数据模型的研究大体上可分为两个方向：一是三维矢量模型，其是用一些基元及其组合去表示三维空间目标，这些基元本身是可以用简单数学解析函数描述的。二是体模型，以体元（Voxel）模型为代表，这种体元模型的特点是易于表达三维空间属性的非均衡变化，其缺点是所占存储空间大、处理时间长。

1．3．3 分布式空间数据模型

分布式空间数据库管理系统和联邦空间数据库是国际上关于分布式空间数据模型的两个主要研究方向。

1）分布式空间数据库管理系统

分布式空间数据库管理系统是将空间数据库技术与计算机网络技术相结合，利用计算机网络对通过通讯线路相关联的空间数据库进行数据和程序的分布处理，以实现集中与分布的统一，即分布式空间数据库管理系统是将分散的空间数据库连成一体。其主要问题包括空间数据的分割、分布式查询、分布式并发控制。

2）联邦空间数据库（Federated Spatial Database）

联邦空间数据库则是在不改变不同来源的各空间数据库管理系统的前提下，将非均质的空间数据库系统联成一体，形成联邦式的空间数据库管理体系，并向用户提供统一的视图。

1．3．4 CASE工具*

CASE 工具是计算机信息系统结构化分析、数据流程描述、数据实体关系表达、数据字典与系统原型生成、原代码生成的重要工具，在非空间型计算机信息系统的设计与建立中有着较为广泛的应用。当前国际上的一个重要发展方向是，根据 GIS 空间数据建模的特点和 CASE 工具的原理，在现有 CASE 软件平台上，发展 GIS 空间数据建模与系统设计的专用功能，这将有效地提高 GIS 空间数据建模及其应用系统设计的自动化程度和技术水平。

2．场模型

对于模拟具有一定空间内连续分布特点的现象来说，基于场的观点是合适的。例如，空气中污染物的集中程度、地表的温度、土壤的湿度水平以及空气与水的流动速度和方向。根据应用的不同，场可以表现为二维或三维。一个二维场就是在二维空间中任何已知的地点上，都有一个表现这一现象的值；而一个三维场就是在三维空间中对于任何位置来说都有一个值。一些现象，诸如空气污染物在空间中本质上讲是三维的，但是许多情况下可以由一个二

---

* 在“地理信息系统软件工程技术”一章较为详尽的描述了该领域的内容。

维场来表示。

场模型可以表示为如下的数学公式：

$z : s \rightarrow z(s)$

上式中，**z** 为可度量的函数，**s** 表示空间中的位置，因此该式表示了从空间域（甚至包括时间坐标）到某个值域的映射。表 3-1 给出了地理研究中一些常模型的例子[A. Vckovski]。

场模型	定义域维数	值域维数	自变量	因变量
$T(z)$	1	1	空间坐标（高程）	高度 z 处的气温
$E(t)$	1	3	时间坐标	某时刻的静电力
$H(x,y)$	2	1	空间坐标	地表高程
$P(x,y,z)$	3	1	空间坐标	土壤的孔隙度
$v(\lambda, \phi, z)$	3	3	空间坐标（λ，φ经纬度，z 高度）	风速（三维矢量）
$\sigma(x, y, z)$	3	9	空间坐标	压力张量
$\Theta(\lambda, \phi, p, t)$	4	1	p 压力面，t 时间	潜温
$\Theta_t(\lambda, \phi, p)$	3	∞	p 压力面	时间序列的潜温
$I(x,y,z,t, \lambda)$	5	1	x, y, z, t 时空坐标，λ 波长	波长 λ 的电磁波在 x, y, z, t 处的辐射强度

2．1 场的特征

场经常被视为由一系列等值线组成，一个等值线就是地面上所有具有相同属性值的点的有序集合。

2．1．1 空间结构特征和属性域

在实际应用中，“空间”经常是指可以进行长度和角度测量的欧几里德空间。空间结构可以是规则的或不规则的，但空间结构的分辨率和位置误差则十分重要，它们应当与空间结构设计所支持的数据类型和分析相适应。属性域的数值可以包含以下几种类型：名称、序数、间隔和比率。属性域的另一个特征是支持空值，如果值未知或不确定则赋予空值。

2．1．2 连续的、可微的、离散的

如果空间域函数连续的话，空间域也就是连续的，即随着空间位置的微小变化，其属性值也将发生微小变化，不会出现像数字高程模型中的悬崖那样的突变值。只有在空间结构和属性域中恰当地定义了“微小变化”，“连续”的意义才确切；

当空间结构是二维（或更多维）时，坡度——或者称为变化率——不仅取决于特殊的位置，而且取决于位置所在区域的方向分布（图 3-2）。连续与可微分两个概念之间有逻辑关系，每个可微函数一定是连续的，但连续函数不一定可微。

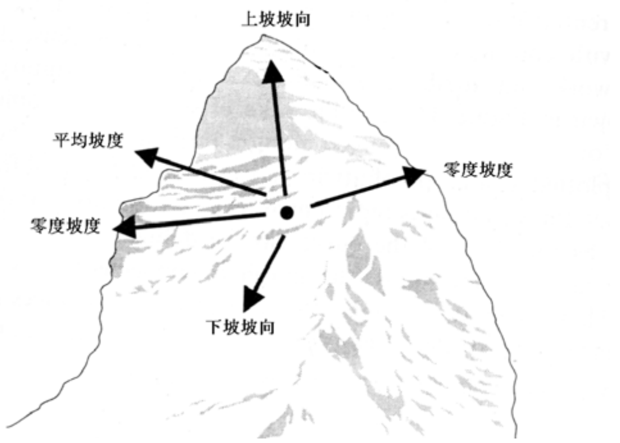
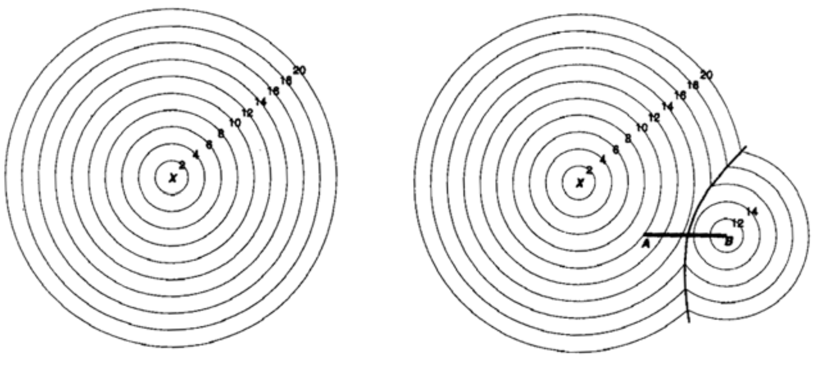


图 3-2：某点的坡度取决于位置所在区域的各方向上的可微性

如果空间域函数是可微分的，空间域就是可微分的；行政区划的边界变化是离散的一个例子，如果目前测得的边界位于 A，而去年这时边界位于 B，但这并不表明 6 个月前边界将位于 BA 之间的中心，边界具有不连续跃变。

2．1．3 与方向无关的和与方向有关的（各向同性和各向异性）

空间场内部的各种性质是否随方向的变化而发生变化，是空间场的一个重要特征。如果一个场中的所有性质都与方向无关，则称之为各向同性场(Isotropic Field)。例如旅行时间，假如从某一个点旅行到另一个点所耗时间只与这两点之间的欧氏几何距离成正比，则从一个固定点出发，旅行一定时间所能到达的点必然是一个等时圆，如图 3-3-(a)所示。如果某一点处有一条高速通道，则利用与不利用高速通道所产生的旅行时间是不同的，见图 3-3-(b)。等时线已标明在图中，图中的双曲线是利用与不利用高速通道的分界线。本例中的旅行时间与目标点与起点的方位有关，这个场称为各向异性场（Anisotropic Field）。



(a) (b)

图 3-3：在各向同性与各向异性场中的旅行时间面

2．1．4 空间自相关

空间自相关是空间场中的数值聚集程度的一种量度。距离近的事物之间的联系性强于距离远的事物之间的联系性。如果一个空间场中的类似的数值有聚集的倾向，则该空间场就表现出很强的正空间自相关；如果类似的属性值在空间上有相互排斥的倾向，则表现为负空间

自相关（图 3-4）。因此空间自相关描述了某一位置上的属性值与相邻位置上的属性值之间的关系。

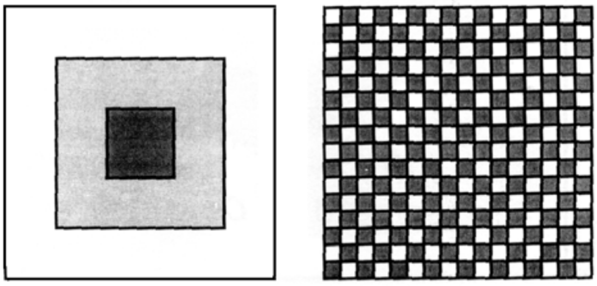


图 3-4：强空间正负自相关模式

2．2 栅格数据模型

栅格数据模型是基于连续铺盖的，它是将连续空间离散化，即用二维铺盖或划分覆盖整个连续空间；铺盖可以分为规则的和不规则*的，后者可当做拓扑多边形处理，如社会经济分区、城市街区；铺盖的特征参数有尺寸、形状、方位和间距。对同一现象，也可能有若干不同尺度、不同聚分性（Aggregation or Subdivisions）的铺盖。在边数从 3 到N的规则铺盖（Regular Tesselations）中，方格、三角形和六角形是空间数据处理中最常用的。三角形是最基本的不可再分的单元，根据角度和边长的不同，可以取不同的形状，方格、三角形和六角形可完整地铺满一个平面（图 3-5）。

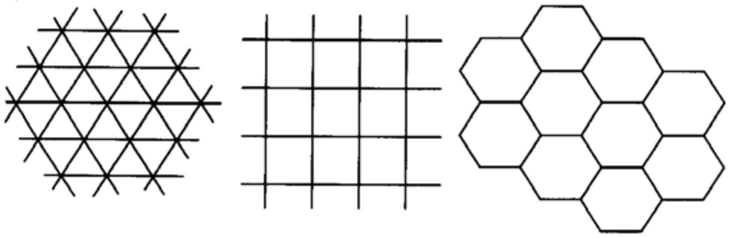


图 3-5：三角形、方格和六角形划分

基于栅格的空间模型把空间看作像元（Pixel）的划分（Tessellation），每个像元都与分类或者标识所包含的现象的一个记录有关。像元与“栅格”两者都是来自图像处理的内容，其中单个的图像可以通过扫描每个栅格产生。GIS 中栅格数据经常是来自人工和卫星遥感扫描设备中，以及用于数字化文件的设备中。采用栅格模型的信息系统，通常应用了前面所述的分层的方法。在每个图层中栅格像元记录了特殊的现象的存在。每个像元的值表明了在已知类中现象的分类情况（图 3-6）。

* “空间分析”一章中提及的Voronoi多边形和TIN属于不规则铺盖。

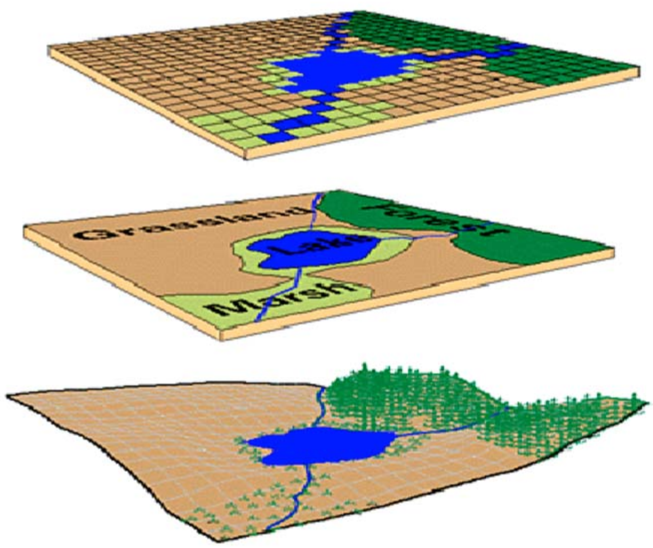


图 3-6：栅格数据模型

由于像元具有固定的尺寸和位置，所以栅格趋向于表现在一个“栅格块”中的自然及人工现象。因此分类之间的界限被迫采用沿着栅格像元的边界线。一个栅格图层中每个像元通常被分为一个单一的类型。这可能造成对现象的分布的误解，其程度则取决与所研究的相关的像元的大小。如果像元针对特征而言是非常小的，栅格可以是一个来表现自然现象的边界随机分布的特别有效的方式，该现象趋于逐渐地彼此结合，而不是简单地划分。如果每个像元限定为一个类，栅格模型就不能充分地表现一些自然现象的转换属性。除非抽样被降低到一个微观的水平，否则许多数据类事实上都是混合类。模糊的特征通过混合像元，在一个栅格内可以被有效地表达，其中组成分类通过像元所有组成度量的或者预测的百分比来表示。尽管如此，也应该强调一个栅格的像元仅仅被赋予一个单一的值。

为了 GIS 数据处理，栅格模型的一个重要的特征就是每个栅格中的像元的位置被预先确定，所以很容易进行重叠运算以比较不同图层中所存储的特征。由于像元位置是预先确定的，且是相同的，在一个具体的应用的不同的图层中，每个属性可以从逻辑上或者从算法上与其它图层中的像元的属性相结合以便产生相应的重叠中一个的属性值。其不同于基于图层的矢量模型之处，在于图层中的面单元彼此是独立的，直接地比较图层必须作进一步处理以识别重叠的属性。

体元（Voxels）：GIS 中基于的栅格表示可以被扩展到三维以产生一个体元（Voxel）模型，其中像元是由长方形，典型是立方体、立体元素所组成。地理数据的一些类型，并不总是由边界表示的，因为数据值可能与一个属性相关，而该属性随着位置的变化而变化，而且并不是清楚地理解边界。这类模型的数据的一个比较合适的模型就是体元模型。该模型被广泛地应用于媒体成像，其中它们源于计算机辅助断层（CT）及核磁反应扫描仪。它们很好地表现渐进的、特殊的位置变化，并适于产生这种变化的剖面图。

3．要素模型

3．1 欧氏（Euclidean）空间和欧氏空间中的三类地物要素

许多地理现象模型建立的基础就是嵌入（Embed）在一个坐标空间中，在这种坐标空间

中，根据常用的公式就可以测量点之间的距离及方向，这个带坐标的空间模型叫做欧氏空间，它把空间特性转换成实数的元组（Tuples）特性，两维的模型叫做欧氏平面。欧氏空间中，最经常使用的参照系统是笛卡尔坐标系（Cartesian Coordinates），它是由一个固定的、特殊的点为原点，一对相互垂直且经过原点的线为坐标轴。此外，在某些情况下，也经常采用其它坐标系统，如极坐标系（Polar Coordinates）。

将地理要素嵌入到欧氏空间中，形成了三类地物要素对象，即点对象、线对象和多边形对象。

3．1．1 点对象

点是有特定的位置，维数为零的物体，包括：

．点实体（Point Entity）：用来代表一个实体；

．注记点：用于定位注记；

．内点（Label Point）：用于记录多边形的属性，存在于多边形内；

．结点（节点）（Node）：表示线的终点和起点；

．角点（Vertex）：表示线段和弧段的内部点。

3．1．2 线对象

线对象是 GIS 中非常常用的维度为 1 的空间组分，表示对象和它们边界的空间属性，由一系列坐标表示，并有如下特征：

．实体长度：从起点到终点的总长；

．弯曲度：用于表示像道路拐弯时弯曲的程度；

．方向性：水流方向是从上游到下游，公路则有单向与双向之分。

线状实体包括线段、边界、链、弧段、网络等，多边线如图 3-7 所示。

3．1．3 多边形对象

面状实体也称为多边形，是对湖泊、岛屿、地块等一类现象的描述。通常在数据库中由一封闭曲线加内点来表示。面状实体有如下空间特性：

．面积范围；

．周长；

．独立性或与其它的地物相邻，如中国及其周边国家；

．内岛或锯齿状外形，如岛屿的海岸线封闭所围成的区域等；

．重叠性与非重叠性，如报纸的销售领域，学校的分区，菜市场的服务范围等都有可能出现交叉重叠现象，一个城市的各个城区一般说来相邻但不会出现重叠。

在计算几何中，定义了许多不同类型的多边形，如图 3-7 所示。

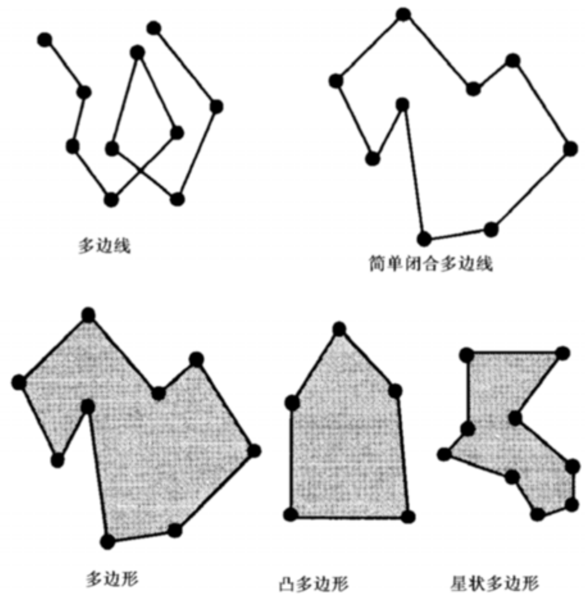


图 3-7：多边线和多边形

3．2 要素模型的基本概念

基于要素的空间模型强调了个体现象，该现象以独立的方式或者以与其它现象之间的关系的方式来研究。任何现象，无论大小，都可以被确定为一个对象（Object），假设它可以从概念上与其邻域现象相分离。要素可以由不同的对象所组成，而且它们可以与其它的相分离的对象有特殊的关系。在一个与土地和财产的拥有者记录有关的应用中，采用的是基于要素的视点，因为每一个土地块和每一个建筑物必须是不同的，而且必须是唯一标识的并且可以单个地测量。一个基于要素的观点是适合于已经组织好的边界现象的，尽管并不被限定。因此，这也适合于人为现象的，例如，建筑物、道路、设施和管理区域。一些自然现象，如湖、河、岛及森林，经常被表现在基于要素的模型中的，因为它们为了某些目的，可以被看成为离散的现象，但应该记住的是，这样现象的边界随着时间的变化很少是固定的，因此，在任何时刻，它们的实际的位置定义很少是精确的。

基于要素的空间信息模型把信息空间分解为对象（Object）或实体（Entity）。一个实体必须符合三个条件：

．可被识别；

．重要（与问题相关）；

．可被描述（有特征）。

而有关实体的特征，可以通过静态属性（如城市名）、动态的行为特征和结构特征来描述实体。与基于场的模型不同，基于要素的模型把信息空间看作许多对象（城市、集镇、村庄、区）的集合，而这些对象又具有自己的属性（如人口密度、质心和边界等）。基于要素的模型中的实体可采用多种维度来定义属性，包括：空间维、时间维、图形维和文本/数字维。

空间对象之所以称为“空间的”，是因为它们存在于“空间”之中，即所谓“嵌入式空间”。空间对象的定义取决于嵌入式空间的结构。常用的嵌入式空间类型有：（1）欧氏空间，它允许在对象之间采用距离和方位的量度，欧氏空间中的对象可以用坐标组的集合来表示；（2）量度空间，它允许在对象之间采用距离量度（但不一定有方向）；（3）拓扑空间，它允许在对象之间进行拓扑关系的描述（不一定有距离和方向）；（4）面向集合的空间，它只采

用一般的基于集合的关系，如包含、合并及相交等。

1）欧氏平面上的空间对象类型

图 3-8 表示了在连续的二维欧氏平面上的一种可能的对象继承等级图。

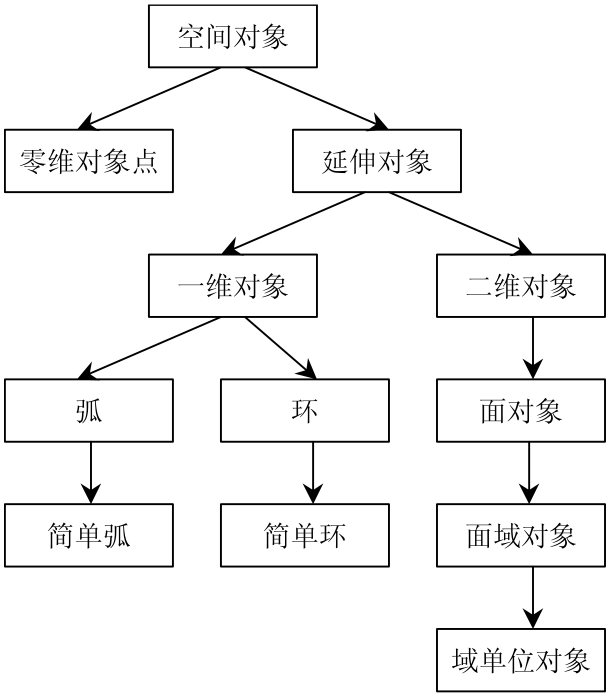


图 3-8：连续空间对象类型的继承等级

在上图中，具有最高抽象层次的对象是“空间对象”类，它派生为零维的点对象和延伸对象，延伸对象又可以派生维一维和二维的对象类。一维对象的两个子类：弧和环（Loop），如果没有相交，则称为简单弧（Simple Arc）和简单环（Simple Loop）。在二维空间对象类中，连通的面对象称为面域对象，没有“洞”的简单面域对象称为域单位对象。

2）离散欧氏平面上的空间对象

欧氏空间的平面因连续而不可计算，必须离散化后才适合于计算。图 3-8 中所有的连续类型的离散形式都存在。图 3-9 表示了部分离散一维对象继承等级关系*。

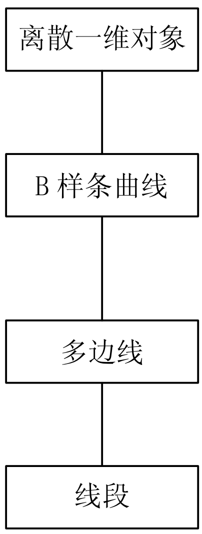


图 3-9：离散一维对象的继承等级

对象行为是由一些操作定义的。这些操作用于一个或多个对象（运算对象），并产生一个新的对象（结果）。可将作用于空间对象的空间操作分为两类：静态的和动态的。静态操作不会导致运算对象发生本质的改变，而动态操作会改变（甚至生成或删除）一个或多个运算对象。

虽然系统的面向对象方法和基于要素的空间数据模型在概念上很相似，但两者之间仍然有着明显的差别。实现基于要素的模型并不一定要求运用面向对象的方法；另一方面，面向对象方法既可以作为描述场的空间模型的框架，也可以作为描述基于要素的空间模型的框架。对于基于要素的模型，采用面向对象的描述是显然合适的；而对于基于场的模型同样可以用面向对象方法来构建。

场和对象可以在多种水平上共存，对于空间数据建模来说，基于场的方法和基于要素的方法并不互相排斥。有些应用可以很自然地应用场来建模，如前面例子中提到的某一区域的气候属性变化就适合于建立场模型；但是，即使是在这种情况下，场模型也并不是适合所有情况。例如，如果采集降雨数据的各个点在空间上很分散且分布无规律，加之这些采集点还有各自的特征，那么，一个包含两个属性，即位置和平均降雨量的对象也许更适合于区域气候属性变化的描述。总之，基于场的模型和基于要素的模型各有长处，应该恰当地综合运用这两种方法来建模。在地理信息系统应用模型的高层建模中、数据结构设计中及地理信息系统应用中，都会遇到这两种模型的集成问题。图 3-10 描述了要素模型和场模型的比较。

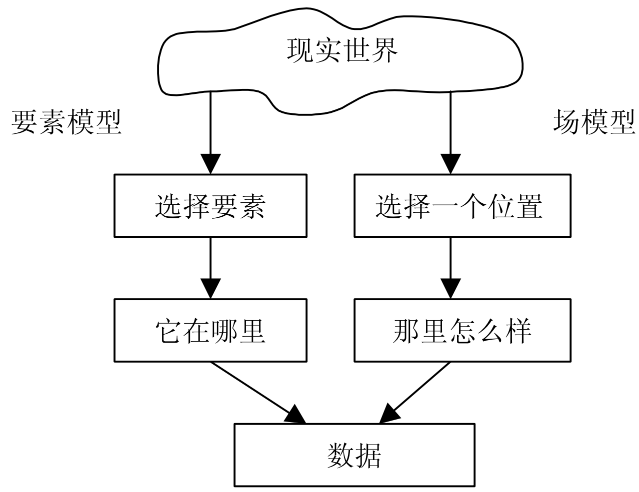


图 3-10：要素模型和场模型的比较[A. Vckovski]

3．3 矢量数据模型

矢量方法（图 3-11）强调了离散现象的存在，由边界线（点、线、面）来确定边界，因此可以看成是基于要素的。然而，在一些基于矢量的 GIS 中，表现表面的便利，带给它模拟二维场的可能性，最常见的例子就是地表高程。栅格技术将重点放置在了空间格网像元位置的内容上，因此经常被描述为基于位置的。栅格数据模型似乎与上面所描述的场的观点相似，但是所储存的空间信息模型并不是对一个连续变量的描述，而它是格网——像元值的一个集合，这些值当然可以被看成抽样一个场模型，但是同样可以被抽样成一个基于对象的模型。

---

* 其中B样条曲线的描述见“空间分析”一章，多边线属于一维的无约束的样条曲线。

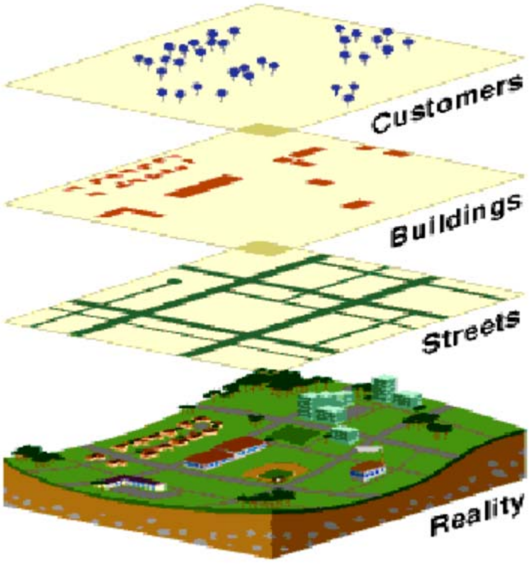


图 3-11：矢量数据模型

矢量数据模型将现象看作原形实体的集合，且组成空间实体。在二维模型内，原型实体是点、线和面；而在三维中，原型也包括表面和体。观察的尺度或者概括的程度，决定了使用的原型的种类。在一个小比例尺表现中，诸如城镇这一现象可以由个别的点所组成，而路和河流由线来表示。当表现的比例尺增大时，必然要考虑到现象的尺度；在一个中等比例尺上，一个城镇可以由特定的原型，如线，来表示用以记录其边界。在较大的比例尺中，城镇将被表现为特定的原型的复杂的集合，包括建筑物的边界、道路、公园以及所包含的其它的自然与管理现象。

矢量模型的表达源于原型空间实体本身，通常以坐标来定义。一个点的位置可以二维或者三维中的坐标的单一集合来描述。一条线通常由有序的两个或者多个坐标对集合来表示。特定坐标之间线的路径可以是一个线性函数或者一个较高次的数学函数，而线本身可以由中间点的集合来确定。一个面通常由一个边界来定义，而边界是由形成一个封闭的环状的一条或多条线所组成。如果区域有个洞在其中，那么可以采用多个环以描述它。

依据应用的类型，对采用矢量数据描述三维模型有一些特殊的要求。地形模型应用要求或者是简单的、单一值的表面（单一值的表面是指对于任意的位置，都有单一的、确定的高程数值），这仅可以表示地表高程；或者它们与地形表面的地形特征相结合，在景观结构中，有必要将地形表面与特征的三维表现结合起来，例如位于其上的建筑物与植被*。为了制图目的，表现地形表面的传统方法可以采用等高线，而对于分析目的而言，等高线并不是一个方便的表示；如果表面被采样为等值线（也许从一个地图上被数字化），它们通常将被转换成最通用的基于GIS的地形表现，如规则格网及不规则三角网。点值的规则格网、或者矩阵，可以直接地来自一个原始的规则的抽样的方案中，通常情况下，是对不规则分布数值的内插，不规则分布数值可以包括数字化等高线和离散点的高程数值。TIN的特征是它们保留了原始的不规则抽样的数据值，它是一个三角化的被用来表现一个三角形的平面，并与这些原始数值相联系。一个TIN单元的表面在缺省情况下被看成平面（Plannar），但是顶点之间也可以采用曲面函数来进行插值。

如果 TIN 被用来表现一个单一值的表面（无论是地形数据还是其它），在与一个插值函数相结合的情况下，它提供了一个二维场的数字化的表现。同样地，如果一个采样点格网同

---

* 与DEM相对应，能够描述地面植被和建筑的模型称为数字表面模型（Digital Surface Model ,DSM），它在建筑规划，生态等领域有着较广泛的用途。

时伴随着一个采样点之间插值函数，它也可以用于实现一个场模型。

如果体对象被存储于基于矢量的 GIS 中，它们通常由闭合的一个或者多个表面来定义；而表面可以由三维线包围的多边形面所定义。线及其构成的点，或顶点的集合，定义了这样的表面为一个多边形网状结构。网的每个表面被视为平面的或者是曲面的；在这两种情况下，需要一个数学的函数用来指示具体坐标之间表面的位置。如果需要一个光滑的表面，可以通过多边形网的顶点来构造数字表面函数，那么，这样一个表面的计算机图形展示就可以通过将数字的表面分解成非常小的平面来实现。数学表面函数的例子如 B 样条函数。这些类型的函数控制了已知的控制点与拟合表面之间的关系，包括了表面的度量以及它与控制点之间的近似程度。

4．基于要素的空间关系分析

4．1 空间关系的基本概念

在地理信息系统中集中存储了以下的内容：

. 空间分布位置信息

. 属性信息

. 拓扑空间关系信息。

由此可见，空间位置、关系与度量的描述在 GIS 中起着举足轻重的作用。

地理要素之间的空间区位关系可抽象为点、线（或弧）、多边形（区域）之间的空间几何关系，其关系如下，如图 3-12 示。

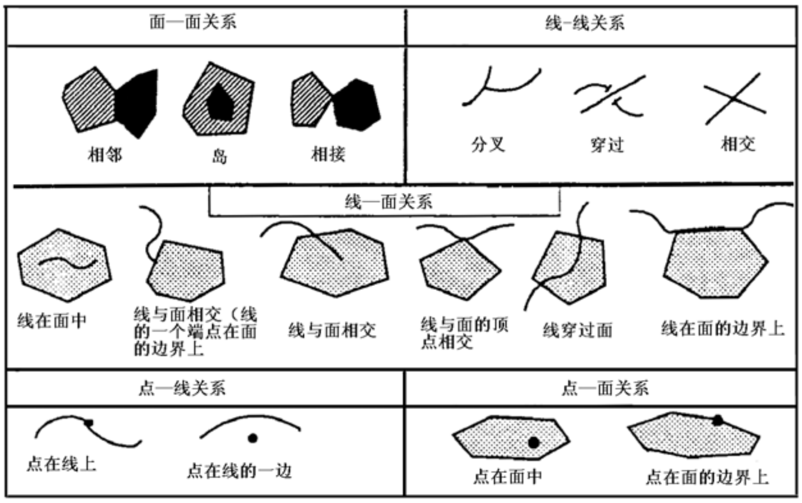


图 3-12：地理要素之间的部分拓扑空间关系

1）点——点关系

相合；

分离；

一点为其它诸点的几何中心；

一点为其它诸点的地理重心。

2）点——线关系

点在线上：可以计算点的性质，如拐点等；

线的端点：起点和终点；

线的交点；

点与线分离：可计算点到线的距离。

3）点——面关系

点在区域内，可以记数和统计；

点为区域的几何中心；

点为区域的地理重心；

点在区域的边界上；

点在区域外部。

4）线——线关系

重合；

相接：首尾环接或顺序相接；

相交：

相切；

并行。

5）线——面关系

区域包含线：可计算区域内线的密度；

线穿过区域：

线环绕区域：对于区域边界，可以搜索其左右区域名称；

线与区域分离。

6）面——面关系

包含：如岛的情形；

相合：

相交：可以划分子区，并计算逻辑与、或、非和异或；

相邻：计算相邻边界的性质和长度；

分离：计算距离、引力等。

近年来，空间关系的理论与应用研究在国内外都非常多。究其原因，一方面是它为地理信息系统数据库的有效建立、空间查询、空间分析、辅助决策等提供了最基本的关系；另一方面是将空间关系理论应用于地理信息系统查询语言，形成一个标准的SQL空间查询语言，可以通过API（Application Program Interface，应用程序接口）进行空间特征的存储、提取、查询、更新等。

空间关系包含三种基本类型，即拓扑关系、方向关系、度量关系。

4．2 拓扑空间关系分析

4．2．1 拓扑属性

拓扑一词来自于希腊文，意思是“形状的研究”。拓扑学是几何学的一个分支，它研究在拓扑变换下能够保持不变的几何属性——拓扑属性。为了得到一些拓扑的感性认识，假设欧氏平面是一张高质量无边界的橡皮，该橡皮能够伸长和缩短到任何理想的程度。想象一下基于这张橡皮所绘制的图形，允许这张纸伸长但是不能撕破或者重叠，这样原来图形的一些属性将保留，而有些属性将会失去。例如，在橡皮表面有一个多边形，多边形内部有一个点。

无论对橡皮进行压缩或拉伸，点依然存在于多边形内部，点和多边形之间的空间位置关系不改变，而多边形的面积则会发生变化。前者则是空间的拓扑属性，后者则不是拓扑属性。表 3-2 列出了包含在欧氏平面中的对象的拓扑和非拓扑属性。

表 3-2：欧氏平面上实体对象所具有的拓扑和非拓扑属性

拓扑属性	一个点在一个弧段的端点 一个弧段是一个简单弧段（弧段自身不相交） 一个点在一个区域的边界上 一个点在一个区域的内部 一个点在一个区域的外部 一个点在一个环的内部 一个面是一个简单面（面上没有“岛”） 一个面的连续性（给定面上任意两点，从一点可以完全在面的内部沿任意路径走向另一点）
非拓扑属性	两点之间的距离 一个点指向另一个点的方向 弧段的长度 一个区域的周长 一个区域的面积

从表中可以看出，拓扑属性描述了两个对象之间的关系，因此又称为拓扑关系（Topological Relation）。图 3-13 为拓扑空间关系的形式化表达。

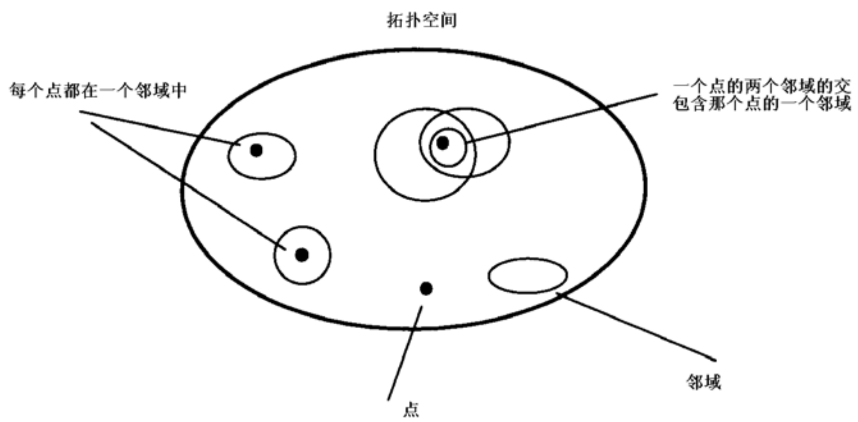


图 3-13：拓扑空间中的点和邻域

4．2．2 拓扑描述的数学基础——点集拓扑学

拓扑学是几何学分支之一，作为近代数学的一门基础理论学科，拓扑学已经渗透到数学的许多分支以及物理、化学和生物学之中，而且在工程技术中也获得了广泛的应用。由于拓扑学是研究图形在拓扑变化下不变的性质，拓扑学已成为地理信息系统空间关系的理论基础，为空间点、线、面之间的包含、覆盖、相离和相接等空间关系的描述提供直接的理论依据。

定义 1：X 为一非空集合，$\rho: X\times X\rightarrow R$ 为一映射，如果对于任意的 x，y，$z\in X$，有:

(1) $\rho(x, y)\geqslant 0$，并且 $\rho(x, y)=0$ 当且仅当 $x=y$；

(2) $\rho(x, y)=\rho(y, x)$；

(3) $\rho(x, z)\leqslant\rho(x, y)+\rho(y, z)$ (三角不等式)；

则称 ρ 为 X 的度量，偶对 (X, ρ) 称为度量空间。

定义 2：A 为度量空间 X 的子集，如果 A 的每一点都有一个球形邻域包含于 A，则称 A 为 ρ 的开集。

定义 3：*X* 为非空集合，*A* 为 *X* 的子集族，如果满足下列条件：

(1) *X* 和空集 *Φ* 属于 *A*；

(2)若 *A1*，*A2* 属于 *A*，则 *A1* 与 *A2* 的交集属于 *A*；

(3)*A* 中任意两个元素的并仍为 A 中的元素；

则称 *A* 为 *X* 的拓扑。如果 *A* 为集合 *X* 的拓扑，则称偶对*(X，A)*为拓扑空间。

定义 4：*A* 为拓扑空间 *X* 的子集。如果点 *x* 属于 *X* 的每一邻域 *U* 中都有 *A* 中异于 *x* 的点，即 *U∩(A～ {x})≠Φ*，则称 *x* 为 *A* 的聚点或极限点。*A* 的聚点可以属于 *A* 也可以不属于 *A*。

定义 5：*A* 为拓扑空间 *X* 的子集，集合 A 的所有聚点构成的集合称为 A 的导集，记作 dX(A)或 d(A)。

定义 6：*A* 为拓扑空间 *X* 的子集，如果 *A* 的每一聚点都属于 *A*，即 *d(A)*为 *A* 的子集，则称 *A* 为闭集。

定义 7：*X* 为拓扑空间，*X* 的子集 *A* 与 *A* 的导集 *d(A)*的并集 *A∪d(A)*称为 *A* 的闭包，记作 *C(A)*。

定义 8：为拓扑空间 *X* 的子集，如果 *A* 是点 *x* 属于 *X* 的邻域，即存在 *X* 的开集 *U* 使得 *x* 属于 *U*，*U* 为 *A* 的子集，则称点 *x* 为集合 *A* 的内点。集合 *A* 的所有内点构成的集合，称为 A 的内部，记作 *I(A)*。

定义 9：*A* 为拓扑空间 *X* 的子集，对于点 *x* 属于 *X*，如果在 *x* 的任一邻域 *U* 中既有 *A* 的点又有～*A* 的点，即：*U∩A≠Φ*并且 *U∩(～A)≠Φ*，则称 *x* 为集合 *A* 的边界点。集合 *A* 的所有边界点的集合称为集合 *A* 的边界，记作 *B(A)*。

定理 1：A 为拓扑空间 X 的任意子集，则:

C(A)=～I(～A)=I(A)∪B(A)

I(A)=～C(～A)=C(A)～B(A)

B(A)=C(A)∩C(～A)=～(I(A)∪I(～A))=B(～A)

4．2．3 拓扑空间关系描述——9 交模型

设有现实世界中的两个简单实体A、B，B(A)、B(B)表示A、B的边界，I(A)、I(B)表示A、B的内部，E(A)、E(B)表示A、B余。Egenhofer[1993]构造出一个由边界、内部、余的点集组成的 9-交空间关系模型(9-Intersection Model,9-IM)如下[*]：

B(A)∩B(B)	*B(A)∩I(B)*	*B(A)∩E(B)*
I(A)∩B(B)	*I(A)∩I(B)*	*I(A)∩E(B)*
E(A)∩B(B)	*E(A)∩I(B)*	*E(A)∩E(B)*

对于该矩阵中的每一元素，都有“空”与“非空”两种取值，9 个元素总共可产生 2^9=512 种情形。

9 交模型形式化地描述了离散空间对象的拓扑关系，基于 9 交模型，可以定义空间数据库的一致性原则，并应用于数据库更新、维护中。此外，9 交模型也是进一步研究空间关系的基础[*]。

9 交模型一共可以表达 512 种可能的空间关系，但是在实际上，有些关系并不存在。表 2 给出了面/面(A/A)，面/线(A/L)，面/点（A/P），线/线（L/L），线/点（L/P），点/点（P/P）可能空间关系的矩阵形式。其中“-”表示不可能存在该关系，“Yb”表示在单值和多值的

---

[*] 在另外一些表述中，B(A)，I(A)，E(A)分别为：∂A，°A和⁻A。

[*] 在这方面，比较有意义的工作是Egenhofer等进行的 9 交模型与自然语言中空间关系描述的对应研究。

矢量图上都可能存在的关系，“Ym”在多值的矢量图上可能存在的关系[*]。

表 3-3：通过 9-交模型表示的两个要素可能的拓扑关系（表中将矩阵的 9 个数值展开得到一个二进制数值，1 表示相应交集不为空，0 表示交集为空）

关系	9-交模型矩阵	A/A	A/L	A/P	L/L	L/P	P/P
r026	000011010	-	-	-	-	-	Yb
r030	000011110	-	-	Yb	-	Yb	-
r031	000011111	Yb	Yb	-	Yb	-	-
r063	000111111	-	Yb	-	Yb	-	-
r092	001011100	-	-	Yb	-	Yb	-
r093	001011101	-	-	-	Yb	-	-
r095	001011111	-	-	-	Yb	-	-
r127	001111111	-	-	-	Yb	-	-
r159	010011111	-	-	-	Yb	-	-
r179	010110011	Ym	-	-	Ym	-	-
r191	010111111	-	Yb	-	Yb	-	-
r220	011011100	Ym	Yb	-	Ym	-	-
r223	011011111	-	-	-	Yb	-	-
r252	011111100	-	Yb	-	-	-	-
r253	011111101	-	Yb	-	-	-	-
r255	011111111	-	Yb	-	Ym	-	-
r272	100010000	-	-	-	-	-	Ym
r277	100010101	-	-	-	Yb	-	-
r279	100010111	Yb	-	-	-	-	-
r284	100011100	-	-	Yb	-	Yb	-
r285	100011101	Yb	Yb	-	-	-	-
r287	100011111	Yb	Yb	-	Yb	-	-
r311	100110111	-	-	-	Yb	-	-
r316	100111100	-	Yb	-	-	-	-
r317	100111101	-	Yb	-	-	-	-
r319	100111111	-	Yb	-	-	-	-
r349	101011101	-	-	-	Yb	-	-
r373	101110101	-	-	-	Yb	-	-
r400	110010000	Ym	-	-	Ym	-	-
r412	110011100	-	Yb	-	-	-	-
r415	110011111	-	-	-	Yb	-	-
r435	110110011	Ym	-	-	Ym	-	-
r439	110110111	-	-	-	Yb	-	-
r444	110111100	-	Yb	-	-	-	-
r445	110111101	-	Yb	-	-	-	-
r447	110111111	-	Yb	-	-	-	-
r476	111011100	Ym	Yb	-	Ym	-	-
r477	111011101	-		-	Yb	-	-
r501	111110101	-		-	Yb	-	-
r508	111111100	-	Yb	-	-	-	-
r509	111111101	-	Yb	-	-	-	-
r511	111111111	Yb	-	-	-	-	-

从上表可以看出，可能的拓扑关系数目要远远少于 512 个（面/面：6，面/线：19，面/点：3，线/线：16，线/点：3，点/点：2）。图 3-14 给出了这些可能关系的图示。从某种意义上讲，9-交模型所描述的拓扑关系只是拓扑关系的类别，对于每一类别可以有多种可能的情形，例如两条相交的线，一个交点的情形和多个交点的 9-交模型表示是一致的，但是其拓扑关系并不同。

[*] 所谓单值矢量图是指对于一给定点，最多只能属于一个地物要素，即地物不能重叠；而对于多值矢量图，地物可以重叠。

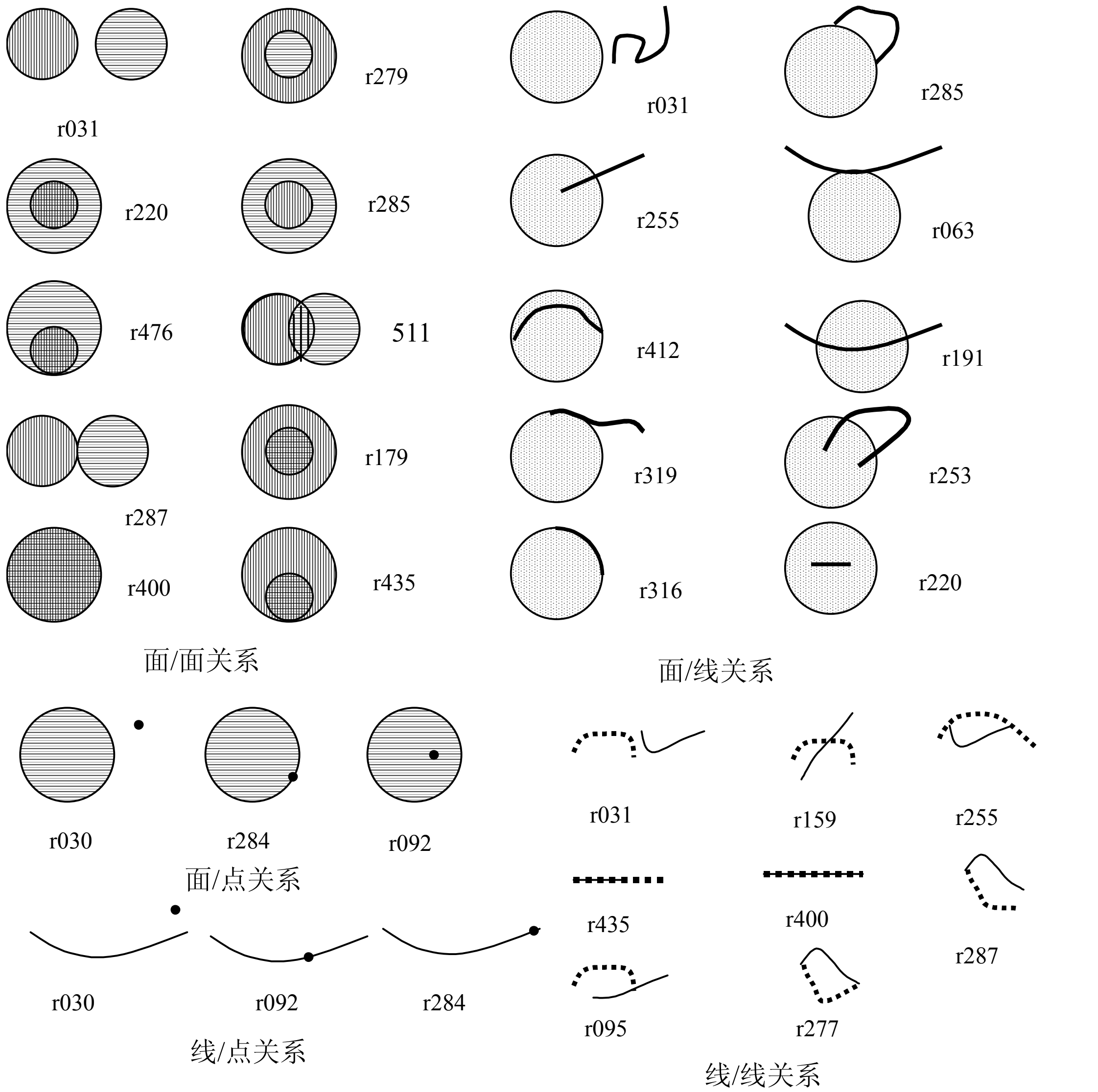


图 3-14：9-交模型所述拓扑关系图示

9 交模型的扩展

通过引进点集的余，9-交空间关系模型增强了面/线、线/线空间关系的唯一性。但它仅仅用“空”与“非空”来区分两个目标的边界、内部、余，对面/面、点/点、点/线、点/面的空间关系描述并无多大改进。为此，该方法仍有一定的局限性。

在地理信息系统中，数据可以划分为几何数据与属性数据两大类型。由于几何数据具有可量测性，为此地理信息系统所涉及的客观世界是一个度量空间，而且每个度量空间又是一个拓扑空间。根据目标的自由度，基本实体可划分为点、线、面三种基本类型。点状目标具有固定的位置和方向，将其定义为零维目标；线状目标都有一条有形或无形的定位线，将其定义为一维目标；面状目标都有一个有形或无形的轮廓线，将其定义为二维目标。

运用维数扩展法，将 9-IM 进行扩展，利用点、线、面的边界、内部、余之间的交集的维数来作为空间关系描述的框架。对于几何实体的边界，它是比其更低一维的几何实体的集合。为此，点的边界为空集；线的边界为线的两个端点，当线为闭曲线时，线的边界为空;面的边界由构成面的所有线构成。若设 P 为一个点集，定义点集的求维函数 DIM 如下：

利用维数扩展法，9 交模型可扩展为

$DIM(B(A)\cap B(B))$	$DIM(B(A)\cap I(B))$	$DIM(B(A)\cap E(B))$
$DIM(I(A)\cap B(B))$	$DIM(I(A)\cap I(B))$	$DIM(I(A)\cap E(B))$

DIM(E(A) ∩B(B))	DIM(E(A) ∩I(B))	DIM(E(A) ∩E(B))

2）根据 DE-9IM，对于点集拓扑空间 X，当需要进行关系判别时，可对矩阵的 9 元取值进行分析、比较。令 C 为各单元交的点集，其取值 P 可能为 *{T，F，*，0，1，2}*。各个取值的具体含义为：

1)*P=T DIM(C)∈ {0，1，2}*，即交集 C 包含有点、线、面；

2)*P=F DIM(C)=-1*，即交集 C 为空；

3)*P=* DIM(C)∈ {-1，0，1，2}*，即两目标交集既有点、线、面，又含有某些部分的交为空的情形，该情况在关系判别时，一般不予以考虑；

4) *P=0 DIM(C)=0*；

5) *P=1 DIM(C)=1*；

6) *P=2 DIM(C)=2*。

扩展 9-交模型中各元素通过取值 *{T，F，*，0，1，2}*，可产生的情形为 6^9=10077696 种，关系非常复杂，通过对大量的空间关系进行归纳和分类，得出 5 种基本的空间关系：相离关系(Disjoint)、相接关系(Touch)、相交关系(Cross)、包含于关系(In)、交叠关系(Overlap)，并将这 5 种关系定义为空间关系的最小集，其特征为：

1) 相互之间不能进行转化；

2) 能覆盖所有的空间关系模式；

3) 能应用于同维与不同维的几何目标；

4) 每一种关系对应于唯一的 DE-9IM 矩阵；

5) 任何其它的 DE-9IM 关系可以通过用这 5 种基本关系进行表达。

4．2．4 拓扑空间关系识别

在地理信息系统中，空间数据具有属性特征、空间特征和时间特征，基本数据类型包括属性数据、几何数据和空间关系数据。作为基本数据类型的空间关系数据主要指点/点、点/线、点/面、线/线、线/面、面/面之间的相互关系。

4．3 方向空间关系分析

4．3．1 方向关系描述

方向关系又称为方位关系、延伸关系，它定义了地物对象之间的方位，如“河北省在河南省北部”就描述了方向关系。为了定义空间目标之间的方向关系，首先定义点目标之间的关系。给定定位参考，即相互垂直的X、Y坐标轴，方向关系的定义采用垂直于坐标轴的直线为参考。令P_i为目标P的点(P为原目标)，Q_j为目标Q的点(Q为参考目标)，$X(P_i)$与$Y(P_i)$函数返回点P_i的X、Y坐标。则P与Q在二维空间中具有以下 8 种可能关系，并提供了一个完整的关系覆盖。这些关系定义为：

$Restricted_East(P_i，Q_j)= X(P_i)>X(Q_j)\ And\ Y(P_i)=Y(Q_j)$

$Restricted_South(P_i，Q_j)=X(P_i)=X(Q_j)\ And\ Y(P_i)<Y(Q_j)$

$Restricted_West(P_i，Q_j)=X(P_i)<X(Q_j)\ And\ Y(P_i)=Y(Q_j)$

$Restricted_North(P_i，Q_j)=X(P_i)=X(Q_j)\ And\ Y(P_i)>Y(Q_j)$

$North_West(P_i，Q_j)=X(P_i)<X(Q_j)\ And\ Y(P_i)>Y(Q_j)$

$North_East(P_i，Q_j)=X(P_i)>X(Q_j)\ And\ Y(P_i)>Y(Q_j)$

$South_West(P_i，Q_j)=X(P_i)<X(Q_j)\ And\ Y(P_i)<Y(Q_j)$

$South_East(P_i，Q_j)=X(P_i)>X(Q_j)\ And\ Y(P_i)<Y(Q_j)$

以上 8 种关系通过点的投影可以精确判断。如有任意两点，上述 8 种关系必有一种满足。而且，这些关系具有传递性，另外一些关系可进行相互转换(如 North_East(PI，QJ)South_West(QI，PJ)，通过对上述 8 种关系进行扩充，可得出另外 4 种方向关系，即：

$East(P_i，Q_j)=North_East(P_i，Q_j)\ Or\ Restricted_East(P_i，Q_j)\ Or\ South_East(P_i，Q_j)$

$South(P_i，Q_j)=South_West(P_i，Q_j)\ Or\ Restricted_South(P_i，Q_j)\ Or\ South_East(P_i，Q_j)$

$West(P_i，Q_j)=North_West(P_i，Q_j)\ Or\ Restricted_West(P_i，Q_j)\ Or\ South_West(P_i，Q_j)$

$North(P_i，Q_j)=North_West(P_i，Q_j)\ Or\ Restricted_North(P_i，Q_j)\ Or\ North_East(P_i，Q_j)$

以点目标之间的方向关系为基础，其余目标之间的方向关系可类似定义。

4．3．2 方向关系识别

MBR(Minimum Bounding Rectangle)指的是空间目标的外切矩形。MBR 的表示非常简单，只需利用两点(左上、右下角点)表示即可。由于 MBR 的简单、实用性，MBR 广泛应用于空间目标数据结构表示以及空间数据查询中。

为了确定目标之间是否具有某种方向关系，首先可判断目标之间的 MBR 是否具有该关系，然后再利用点/点关系进一步进行关系判断，确定具体的关系。

4．4 度量空间关系分析

基本空间对象度量关系包含点/点、点/线、点/面、线/线、线/面、面/面之间的距离。在基本目标之间关系的基础上，可构造出点群、线群、面群之间的度量关系。例如，在已知点/线拓扑关系与点/点度量关系的基础上，可求出点/点间的最短路径、最优路径、服务范围等；已知点、线、面度量关系，进行距离量算、邻近分析、聚类分析、缓冲区分析、泰森多边形分析等。

4．4．1 空间指标量算

定量量测区域空间指标和区域地理景观间的空间关系是地理信息系统特有的能力。其中区域空间指标包括：

1）几何指标：位置、长度（距离）、面积、体积、形状、方位等指标；

2）自然地理参数：坡度、坡向、地表辐照度、地形起伏度、河网密度、切割程度、通达性等；

3）人文地理指标：如集中指标、区位商、差异指数、地理关联系数、吸引范围、交通便利程度、人口密度等。

4．4．2 地理空间的距离度量

地理空间中两点间的距离度量可以沿着实际的地球表面进行，也可以沿着地球椭球体的距离量算，具体的，距离可以表现为以下几种形式（以地球上两个城市之间的距离为例）（图

3-15）：

1）大地测量距离：该距离即沿着地球大圆经过两个城市中心的距离。

2）曼哈顿距离：纬度差加上经度差（名字“曼哈顿距离”是由于在曼哈顿，街道的格局可以被模拟成两个垂直方向的直线的一个集合）。

3）旅行时间距离：从一个城市到另一个城市的最短的时间可以用一系列指定的航线来表示（假设每个城市至少有一个飞机场）。

4）词典编纂距离：在一个固定的地名册中一系列城市中它们位置之间的绝对差值。

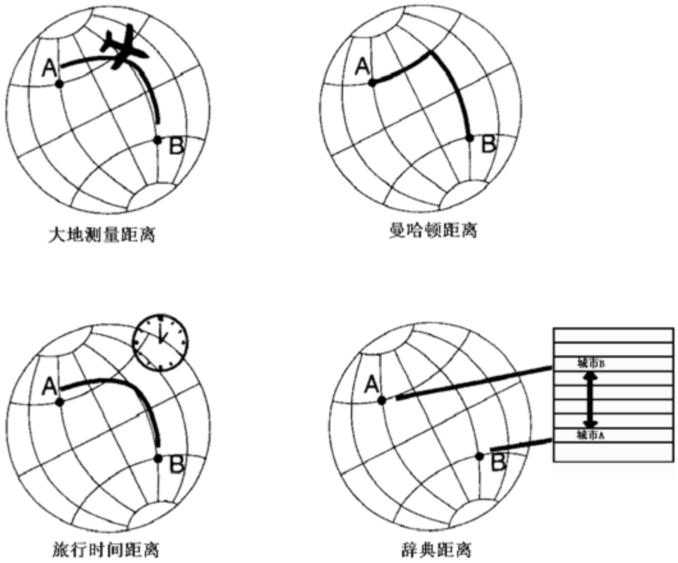
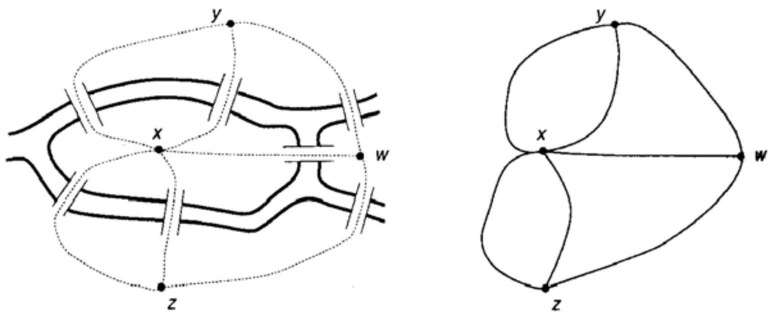


图 3-15：地球上各种形式的距离

5．网络结构模型

5．1 网络空间

网络拓扑系统研究的创始人被公认为数学家 Leonard Euler，他在 1736 年解决了当时一个著名的问题，叫做 Konigsberg 桥问题。图 3-16-a 显示了该桥的一个概略的路线图。该问题就是找到一个循环的路，该路只穿过其中每个桥一次，最后返回到起点。一些实验表明这项任务是不可能的，然而，从认为没有这样的路线到说明它的步骤并不是这样容易的。



(a) (b)

图 3-16：Konigsberg Park 中的图形理论模型

Euler 成功地证明了这是一项不可能的任务；或者，换一句话来说，这个问题是没有解的。为了做这件事，他建立了该桥的一个空间模型，该模型抽象出了所有的仅有的桥之间的拓扑关系，见图 3-16-b。实心圆表示结点或顶点。它们被标上 w、x、y、z，并且抽象为陆地面。线表示弧段或边线，它们抽象为陆地之间的直线，并且在每种情况下需要使用一个桥，完整的模型叫做网络或图形。Euler 证明了不可能从一个结点开始，沿着图形的边界，遍历每个边界只有一次，最后到达第一个结点。他所采用的论点是非常简单的，依据的是经过每个结点的边的奇数/偶数。我们看到：除了开始的结点和末端的结点外，经过一个结点的路径必须是沿着一个边界进入，又沿着另一个边界出去。这说明了两个边界对应着那个结点。因此，如每个中间结点相连的边界的数量必须是偶数，当然，如果这个问题是有解的话。图 3-15 中，没有一个结点的边界数是偶数的。因此，这个图论的问题是没有解的，并且与 Konigsberg 桥问题有关的最初的问题也是无解的。

5．2 网络模型

在网络模型中，地物被抽象为链、节点等对象，同时要关注其间连通关系*。基于网络的空间模型与基于要素的模型在一些方面有共同点，因为它们经常处理离散的地物，但是最基本的特征就是需要多个要素之间的影响和交互，通常沿着与它们相连接的通道。相关的现象的精确形状并不是非常重要的，重要的是具体现象之间距离或者阻力的度量。网络模型的典型的例子就是研究交通，包括陆上、海上及航空线路，以及通过管线与隧道分析水、汽油及电力的流动。例如，一个电力供应公司对它们的设施管理可能既采用了一个基于要素的视点，同时又采用了一个基于网络的视点，这依赖于他们关心的是否是替换一个特定的管道，在这种情况下，一个基于要素的视点可能是合适的；或者例如他们关心的是分析重建线路的目的，在这种情况下，网络模型将是合适的。

网状模型的基本特征是，结点数据间没有明确的从属关系，一个结点可与其它多个结点建立联系。网状模型将数据组织成有向图结构。结构中结点代表数据记录，连线描述不同结点数据间的关系。有向图（Digraph）的形式化定义为：

Digraph = (Vertex,{Relation})

其中 Vertex 为图中数据元素（顶点）的有限非空集合；Relation 是两个顶点（Vertex）之间的关系的集合。

有向图结构比树结构具有更大的灵活性和更强的数据建模能力。网状模型可以表示多对多的关系，其数据存储效率高于层次模型，但其结构的复杂性限制了它在空间数据库中的应用。

网状模型反映了现实世界中常见的多对多关系，在一定程度上支持数据的重构，具有一定的数据独立性和共享特性，并且运行效率较高。但它在应用时也存在以下问题：

. 网状结构的复杂，增加了用户查询和定位的困难。它要求用户熟悉数据的逻辑结构，知道自身所处的位置。

. 网状数据操作命令具有过程式性质。

. 不直接支持对于层次结构的表达。

. 基本不具备演绎功能。

. 基本不具备操作代数基础。

---

* 在“空间分析”一章中的“网络分析”部分更加具体的描述了网络模型。

6．时空模型

6．1 时空数据模型概述

6．1．1 研究概述

传统的地理信息系统应用只涉及地理信息的两个方面：空间维度和属性维度，因此也叫SGIS(Static GIS)，而能够同时处理时间维度的 GIS 叫 TGIS（Temporal GIS）。

在 GIS 中，具有时间维度的数据可以分为两类，一类是可以称为结构化的数据，如一个测站历史数据的积累，它可以通过在属性数据表记录中简单地增加一个时间戳（Time Stamp）实现其管理；另一类是非结构化的，最典型的例子是土地利用状况的变化（图 3-17），描述这种数据，是 TGIS 数据模型的重点要解决的问题。

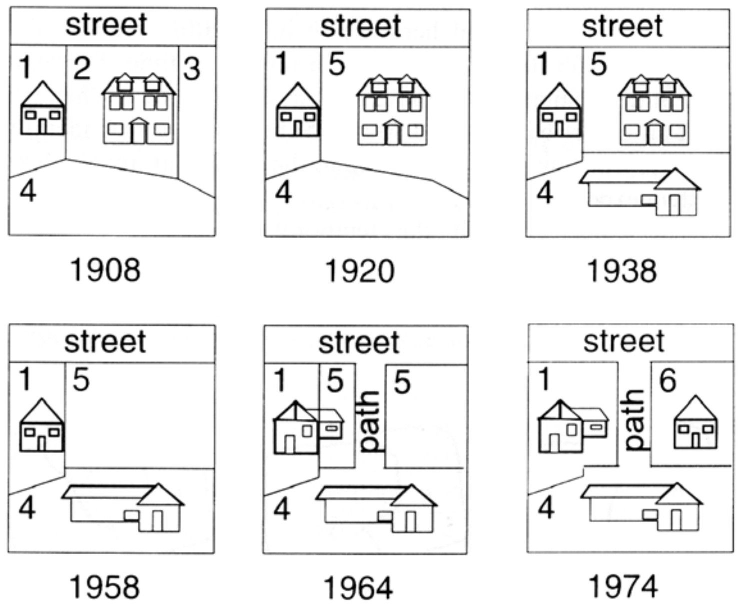


图 3-17：土地利用随时间的推移而变化

TGIS 数据模型特点是语义更丰富、对现实世界的描述更准确，其物理实现的最大困难在于海量数据的组织和存取。TGIS 技术的本质特点是“时空效率”。当前主要的 TGIS 模型包括：空间时间立方体模型（Space-time Cube）；序列快照模型（Sequent Snapshots）；基图修正模型（Base State with Amendments）；空间时间组合体模型（Space-time Composite）。其中序列快照模型和 GIS 分类中的模拟 GIS（Analog GIS）一样，只是一种概念上的模型，不具备实用的开发价值，而其它几种模型都有自己的特点和适用范围，如基图修正模型比较适合于栅格模型的 TGIS 开发。

6．1．2 TGIS 的研究思路

TGIS 海量数据的处理必然导致数学模型的根本变化。TGIS 问题的最终解决在于“可与拓扑论相类比的”全新数学思路的出现。目前可以研究 TGIS 技术，以便在 SGIS 的框架中

用 TGIS 技术实现 TGIS 功能。对 TGIS 模型的研究可以本着两种思路进行平行探索：综合模型和分解模型。先用分解模型思路针对典型应用领域（如土地利用动态监测工作）进行全面研究，同时不断丰富、充实综合模型，最后得到一个比较完善的综合模型。

6．2 时空数据模型设计的基本思想

地籍变更、海岸线变化、土地城市化、道路改线、环境变化等应用领域，需要保存并有效地管理历史变化数据，以便将来重建历史状态、跟踪变化、预测未来。这就要求有一个组织、管理、操作时空数据的高效时空数据模型。时空数据模型是一种有效组织和管理时态地理数据，属性、空间和时间语义更完整的地理数据模型。

一个合理的时空数据模型必须考虑以下几方面的因素：节省存储空间、加快存取速度、表现时空语义。时空语义包括地理实体的空间结构、有效时间结构、空间关系、时态关系、地理事件、时空关系。时空数据模型设计的基本指导思想：

（1）根据应用领域的特点（如宏观变化观测与微观变化观测）和客观现实变化规律（同步变化与异步变化、频繁变化与缓慢变化），折中考虑时空数据的空间/属性内聚性和时态内聚性的强度，选择时间标记的对象。对于属性，有属性数据项时间标记、实体时间标记、数据库时间标记；对于空间，有坐标点时间标记、弧段时间标记、实体时间标记、数据库时间标记。

（2）同时提供静态（变化不活跃）、动态（变化活跃）数据建模手段（静态、动态数据类型和操作）。当前、历史等不同使用频率的数据分别组织存放，以便存取。一般地，将当前数据存放在本地机磁盘上，而将历史数据存放在远程服务器大容量光盘上。

（3）数据结构里显式表达两种地理事件：地理实体进化事件和地理实体存亡事件。地理事件以事件发生的相关源状态和终止状态表达。构成地理实体存亡事件的源状态有参加事件的实体标识集合表示。时间的本质为事件发生的序列，地理事件序列直接表明地理时间语义。常见的状态变化查询即地理事件查询。

（4）时空拓扑关系一般指地理实体空间拓扑关系的拓扑事件间的时态关系。时空拓扑关系揭示了地理实体在时间和空间上的相关性。为了有效地表达时空拓扑关系，需要存储空间拓扑关系的时变序列。

7．三维模型

目前随着计算机技术的飞速发展和计算机图形学理论的日趋完善，GIS 作为一门新兴的边缘学科也日趋成熟，许多商品化的 GIS 软件功能日趋完善。但是，绝大多数的商品化 GIS 软件包还只是在二维平面的基础上模拟并处理现实世界上所遇到的现象和问题，而一旦涉及到处理三维问题时，往往感到力不从心，GIS 处理的与地球有关的数据，即通常所说的空间数据，从本质上说是三维连续分布的。从事关于地质、地球物理、气象、水文、采矿、地下水、灾害、污染等方面的自然现象是三维的，当这些领域的科学家试图以二维系统来描述它们时，就不能够精确地反映、分析或显示有关信息。三维 GIS 的要求与二维 GIS 相似，但在数据采集、系统维护和界面设计等方面比二维 GIS 要复杂得多。

7．1 三维 GIS 的功能

目前，三维 GIS 所研究的内容以及实现的功能主要包括：

1）数据编码：是采集三维数据和对其进行有效性检查的工具，有效性检查将随着数据的自然属性、表示方法和精度水平的不同而不同。

2）数据的组织和重构：这包括对三维数据的拓扑描述以及一种表示法到另一种表示法的转换（如从矢量的边界表示转换为栅格的八叉树表示）。

3）变换：既能对所有物体或某一类物体，又能对某个物体进行平移、旋转、剪裁、比例缩放等变换。另外还可以将一个物体分解成几个以及将几个物体组合成一个。

4）查询：此功能依赖于单个物体的内在性质（如位置、形状、组成）和不同物体间的关系（如连接、相交、形状相似或构成相似）。

5）逻辑运算：通过与、或、非及异或运算符对物体进行组合运算。

6）计算：计算物体的体积、表面积、中心、物体之间的距离及交角等。

7）分析：如计算某一类地物的分布趋势，或其它指标，以及进行模型的比较。

8）建立模型。

9）视觉变换：在用户选择的任何视点，以用户确定的视角、比例因子、符号来表示所有地物或某些指定物体。

10）系统维护：包括数据的自动备份、安全性措施、以及网络工作管理。

7．2 三维数据结构

三维数据结构同二维一样，也存在栅格和矢量两种形式。栅格结构使用空间索引系统，它包括将地理实体的三维空间分成细小的单元，称之为体元或体元素。存储这种数据的最简单形式是采用三维行程编码，它是二维行程编码在三维空间的扩充。这种编码方法可能需要大量的存储空间，更为复杂的技术是八叉树，它是二维的四叉树的延伸。三维矢量数据结构表示有多种方法，其中运用最普遍的是具有拓扑关系的三维边界表示法和八叉树表示法。

7．2．1 八叉树三维数据结构

用八叉树来表示三维形体，既可以看成是四叉树方法在三维空间的推广，也可以是用三维体素列阵表示形体方法的一种改进。八叉树的逻辑结构如下：假设要表示的形体 V 可以放在一个充分大的正方体 C 内，C 的边长为 2 的 n 次方，形体 VC，它的八叉树可以用以下的递归方法来定义：八叉树的每个节点与 C 的一个子立方体对应，树根与 C 本身相对应，如果 V=C，那么 V 的八叉树仅有树根，如果 V 不等于 C，则 C 等分为八个子立方体，每个子立方体与树根的一个子节点相对应。只要某个子立方体不是完全空白或完全为 V 所占据，就要被八等分，从而对应的节点也就有了八个子节点。这样的递归判断、分割一直要进行到结点所对应的立方体或是完全空白，或者是完全为 V 占据，或是其大小已是预先定义的体素大小，并且对它与 V 之交作一定的“舍入”，使体素或认为是空白的，或认为是 V 占据的。

如此所生成的八叉树上的节点可分为三类：

1）灰节点，对应的立方体部分地为 V 所占据；

2）白节点，所对应的立方体中无 V 的内容；

3）黑节点，所对应的立方体全为 V 所占据。

后两类又称为叶结点。由于八叉树的结构与四叉树的结构是非常相似的，所以八叉树的存储结构方式可以完全沿用四叉树的有关方法。根据不同的存储方式，八叉树也可以分别称为常规的、线形的、一对八的八叉树等等。

1）规则的八叉树

八叉树的存储结构是用一个有九个字段的记录来表示树中的每个结点，其中一个字段用来描述该结点的特性，其余的八段用来作为存放指向其八个子结点的指针。这是最普通使用的表示树形数据的存储结构方式。规则八叉树缺陷较多，最大的问题是指针占用了大量的空间。因此，这种方式虽然十分自然，容易掌握，但在存储空间的使用率方面不很理想。

2）线形八叉树

线形八叉树注重考虑如何提高空间利用率，用某一预先确定的次序遍历八叉树，将八叉树转换成一贯线形表，表的每个元素与一个结点相对应。线形八叉树不仅节省存储空间，对某些运算也较为方便。但是为此付出的代价是丧失了一定的灵活性，如图 3-18 和图 3-19 所示。

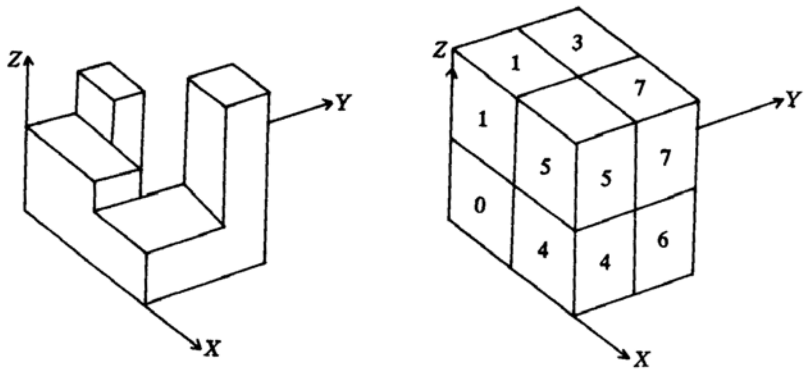


图 3-18：体元形式的三维数据

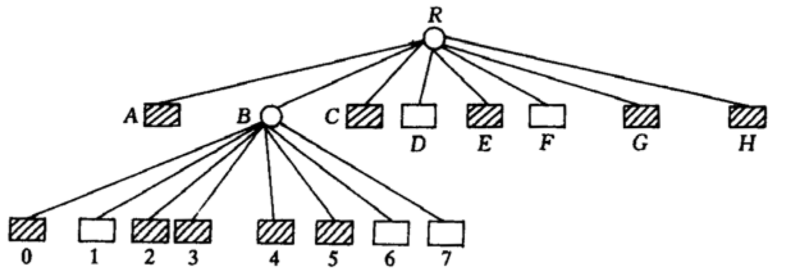


图 3-19：图 3-18 的线性八叉树编码

3）一对八式的八叉树

一个非叶结点有八个子节点，为了确定起见，将它们分别标记为 0，1，2，3，4，5，6，7。从上面的介绍可以看到，如果一个记录与一个结点相对应，那么在这个记录中描述的是这个结点的八个子结点的特征值。而指针给出的则是该八个子结点所对应记录的存放处，而且还隐含地假设了这些子结点记录存放的次序。也就是说，即使某个记录是不必要的，那么相应的存储位置也必须空闲在那里，以保证不会错误地存取到其它同辈结点的记录。这样当然会有一定的浪费，除非它是完全的八叉树，即所有的叶结点均在同一层次出现，而在该层次之上的所有层中的结点均为非结点。为了克服这种缺陷，一是增加计算量，即在存取相应结点记录之前，首先检查它的父结点记录，看一下之前有几个叶结点，从而可以知道应该如何存取所需结点记录。这种方法的存储需求无疑是最小的，但是要增加计算量；另一个是在记录中增加一定的信息，使计算工作适当减少或者更方便。例如在原记录中增加三个字节，一分为八，每个子结点对应三位，代表它的子结点在指针指向区域中的偏移。因此，要找到

它的子结点的记录位置，只要固定地把指针指向的位置加上这个偏移值（0-7）乘上记录所占的字节数，就是所要的记录位置，因而一个结点的描述记录为：

偏移	指针	SWB	SWT	NWB	NWT	SEB	SET	NEB	NET

用这种方式所得到的八叉树和以前相同，只是每个记录前多了三个字节。

7．2．2 三维数据的显示

三维显示通常采用截面图、等距平面、多层平面和立体块状图等多种表现形式，大多数三维显示技术局限于 CRT 屏幕和绘图纸的二维表现形式，人们可以观察到地理现象的三维形状，但不能将它们作为离散的实体进行分析，如立体不能被测量、拉伸、改变形状或组合。借助三维显示技术，通过离散的高程点形成等高线图、截面图、多层平面和透视图，可以把这些最初都是人工完成的工作，用各种计算机程序迅速高效地完成。图 3-20 给出了一种三维数据的表示方法。

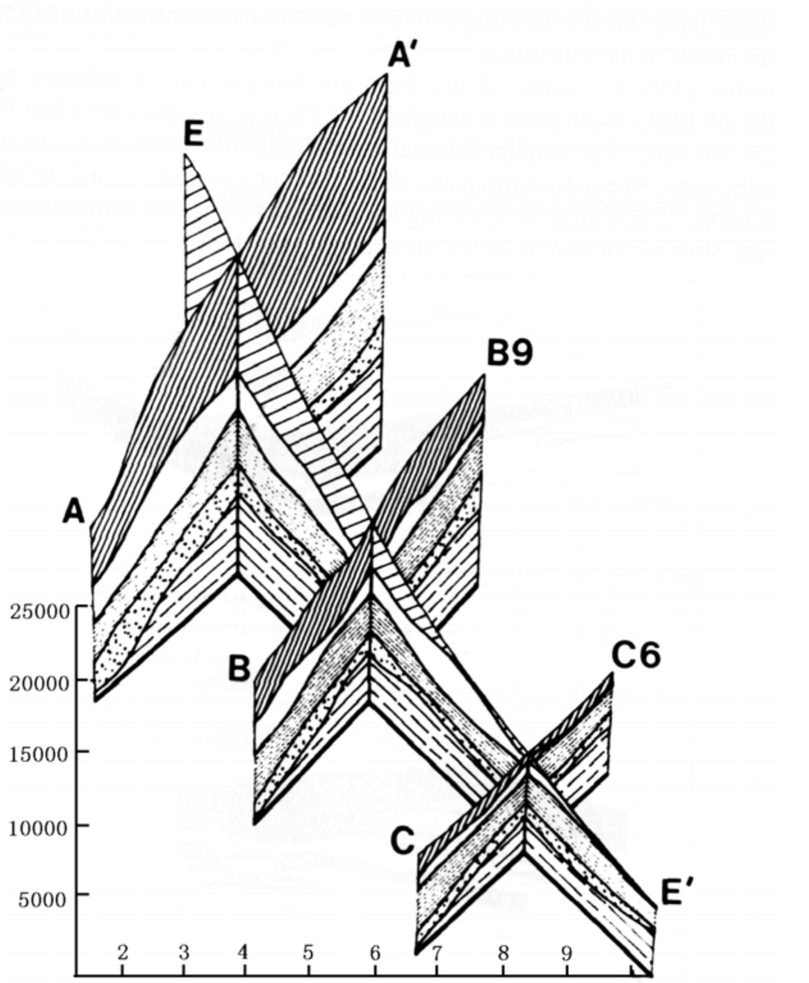


图 3-20：通过“围墙”状的剖面表示三维数据

如今我在什么地方?我不知道那是什么地方。我猜不着。到底这里是哪里?映入我眼帘的只是不知何处去的人蔓，行色匆匆地从我身边走过去。

村上春树

第四章 空间参照系统和地图投影

导读：正如上一章所描述的，一个要素要进行定位，必须嵌入到一个空间参照系中，因为GIS所描述是位于地球表面的信息，所以根据地球椭球体建立的地理坐标（经纬网）可以作为所有要素的参照系统。因为地球是一个不规则的球体，为了能够将其表面的内容显示在平面的显示器或纸面上，必须进行坐标变换。

本章讲述了地球椭球体参数、常见的投影类型。考虑到目前使用的1：100万以上地形图都是采用高斯——克吕格投影，本章最后又对该种投影类型和相关的地形图分幅标准做了简单介绍。

1．地球椭球体基本要素

1．1 地球椭球体

1．1．1 地球的形状

为了从数学上定义地球，必须建立一个地球表面的几何模型。这个模型由地球的形状决定的。它是一个较为接近地球形状的几何模型，即椭球体，是由一个椭圆绕着其短轴旋转而成。

地球自然表面是一个起伏不平、十分不规则的表面，有高山、丘陵和平原，又有江河湖海。地球表面约有71%的面积为海洋所占用，29%的面积是大陆与岛屿。陆地上最高点与海洋中最深处相差近20公里。这个高低不平的表面无法用数学公式表达，也无法进行运算。所以在量测与制图时，必须找一个规则的曲面来代替地球的自然表面。当海洋静止时，它的自由水面必定与该面上各点的重力方向（铅垂线方向）成正交，我们把这个面叫做水准面。但水准面有无数多个，其中有一个与静止的平均海水面相重合。可以设想这个静止的平均海水面穿过大陆和岛屿形成一个闭合的曲面，这就是大地水准面（图4-1）。

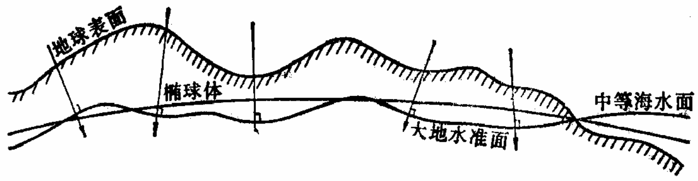


图4-1：大地水准面

大地水准面所包围的形体，叫大地球体。由于地球体内部质量分布的不均匀，引起重力方向的变化，导致处处和重力方向成正交的大地水准面成为一个不规则的，仍然是不能用数学表达的曲面。大地水准面形状虽然十分复杂，但从整体来看，起伏是微小的。它是一个很接近于绕自转轴（短轴）旋转的椭球体。所以在测量和制图中就用旋转椭球来代替大地球体，

这个旋转球体通常称地球椭球体，简称椭球体。

1．1．2 地球的大小

关于地球椭球体的大小，由于采用不同的资料推算，椭球体的元素值是不同的。现将世界各国常用的地球椭球体的数据列表如下：

表 4-1：各种地球椭球体模型

椭球体名称	年代	长半轴（米）	短半轴（米）	扁率
白塞尔(Bessel)	1841	6377397	6356079	1：299.15
克拉克(Clarke)	1880	6378249	6356515	1：293.5
克拉克(Clarke)	1866	6378206	6356584	1：295.0
海福特(Hayford)	1910	6378388	6356912	1：297
克拉索夫斯基	1940	6378245	6356863	1：298.3
I．U．G．G	1967	6378160	6356775	1：298.25
埃维尔斯特(Everest)	1830	6377276	6356075	1：300.8

1．1．3 椭球体的半径

地球椭球体表面是一个规则的数学表面。椭球体的大小，通常用两个半径：长半径 a 和短半径 b，或由一个半径和扁率来决定。扁率 α 表示椭球的扁平程度。扁率的计算公式为：$\alpha=(a-b)/a$。这些地球椭球体的基本元素 a、b、α 等，由于推求它的年代、使用的方法以及测定的地区不同，其结果并不一致，故地球椭球体的参数值有很多种。中国在 1952 年以前采用海福特（Hayford）椭球体，从 1953-1980 年采用克拉索夫斯基椭球体。随着人造地球卫星的发射，有了更精密的测算地球形体的条件。1975 年第 16 届国际大地测量及地球物理联合会上通过国际大地测量协会第一号决议中公布的地球椭球体，称为 GRS（1975），中国自 1980 年开始采用 GRS（1975）新参考椭球体系。由于地球椭球长半径与短半径的差值很小，所以当制作小比例尺地图时，往往把它当作球体看待，这个球体的半径为 6371 公里。

1．1．4 高程

地面点到大地水准面的高程，称为绝对高程。如图 2 所示，P_0P_0'为大地水准面，地面点A和B到P_0P_0'的垂直距离H_A和H_B为A、B两点的绝对高程。地面点到任一水准面的高程，称为相对高程。如图 2 中，A、B两点至任一水准面P_1P_1'的垂直距离H_A'和H_B'为A、B两点的相对高程。

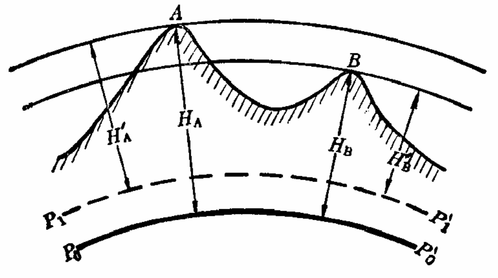


图 4-2：地面点的高程

我国的大地控制网

我国面积辽阔，在约 960 万平方公里的土地上进行测图工作，需要分成若干单元测区，而且测量的精度又要符合统一要求，为此，在全国范围内建立统一的大地控制网。控制网分为平面控制网和高程控制网。

大地坐标：在地面上建立一系列相连接的三角形，量取一段精确的距离作为起算边，在这个边的两端点，采用天文观测的方法确定其点位（经度、纬度和方位角），用精密测角仪器测定各三角形的角值，根据起算边的边长和点位，就可以推算出其他各点的坐标。这样推算出的坐标，称为大地坐标。

我国 1954 年在北京设立了大地坐标原点，由此计算出来的各大地控制点的坐标，称为 1954 年北京坐标系。我国 1986 年宣布在陕西省泾阳县设立了新的大地坐标原点，并采用 1975 年国际大地测量协会推荐的大地参考椭球体，由此计算出来的各大地控制点坐标，称为 1980 年大地坐标系。

我国高程的起算面是黄海平均海水面。1956 年在青岛设立了水准原点，其他各控制点的绝对高程都是根据青岛水准原点推算的，称此为 1956 年黄海高程系。1987 年国家测绘局公布：中国的高程基准面启用《1985 国家高程基准》取代国务院 1959 年批准启用的《黄海平均海水面》。《1985 国家高程基准》比《黄海平均海水面》上升 29 毫米。

1．2 地图比例尺

1．2．1 比例尺表示法

地图比例尺通常认为是地图上距离与地面上相应距离之比。地图比例尺可用下述方法表示。

1）数字比例尺

这是简单的分数或比例，可表示为 1：1000000 或 1/1000000，最好用前者。这意味着，地图上（沿特定线）长度 1 毫米、1 厘米或 1 英寸（分子），代表地球表面上的 1000000 毫米、厘米或英寸（分母）。

2）文字比例尺

这是图上距离与实地距离之间关系的描述。例如，1：1000000 这一数字比例尺可描述为“图 1 毫米等于实地 1 公里”。

3）图解比例尺或直线比例尺

这是在地图上绘出的直线段，常常绘于图例方框中或图廓下方，表示图上长度相当于实地距离的单位。

4）面积比例尺

这关系到图上面积与实地面积之比，表示图上 1 单位面积（平方厘米）与实地上同一种平方单位的特定数量之比。

1．2．2 比例系数

表明确定的比例尺与实际比例尺数值之间的关系叫做比例系数（SF）。可以这样理解比例系数，首先将地球缩小为所选比例尺的地球仪地图；然后将该球形地图转换为平面地图。上述平面地图的数字比例尺就是地球仪的比例尺，叫做主比例尺（或名义比例尺）；真实比例尺就是平面地图上的实际比例尺，当然各处是不相同的。

比例系数可按下式计算：SF=实际比例尺/主比例尺
该公式表明，比例系数是实际比例尺与单位（1）主比例尺之比。当比例系数为 2 时，实际比例尺为主比例尺的两倍。比例系数只在小比例尺世界地图上比较明显。在大比例尺地图上，各处的比例系数对于 1 只有很小的变化。

2．坐标系

所谓坐标系，包含两方面的内容：一是在把大地水准面上的测量成果化算到椭球体面上的计算工作中，所采用的椭球的大小；二是椭球体与大地水准面的相关位置不同，对同一点的地理坐标所计算的结果将有不同的值。因此，选定了一个一定大小的椭球体，并确定了它与大地水准面的相关位置，就确定了一个坐标系（图 4-3）。

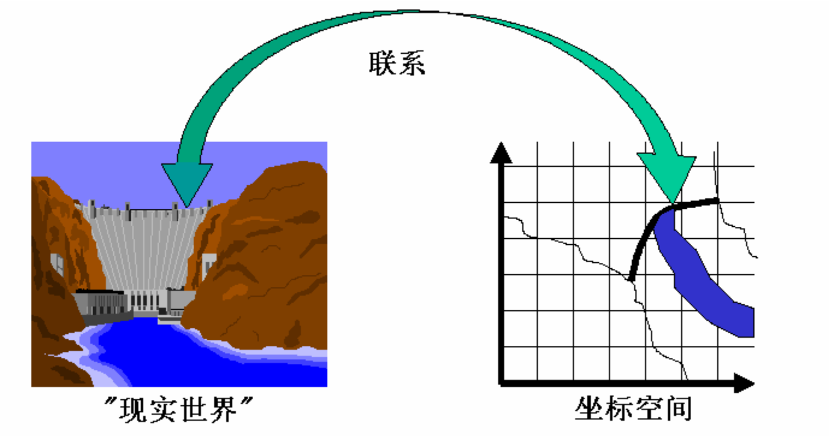


图 4-3：现实世界和坐标空间的联系

2．1 地理坐标

地球除了绕太阳公转外，还绕着自己的轴线旋转，地球自转轴线与地球椭球体的短轴相重合，并与地面相交于两点，这两点就是地球的两极，北极和南极。垂直于地轴，并通过地心的平面叫赤道平面，赤道平面与地球表面相交的大圆圈（交线）叫赤道。平行于赤道的各个圆圈叫纬圈（纬线）（Parallel），显然赤道是最大的一个纬圈。

通过地轴垂直于赤道面的平面叫做经面或子午圈（Meridian），所有的子午圈长度彼此都相等。（图 4-4）

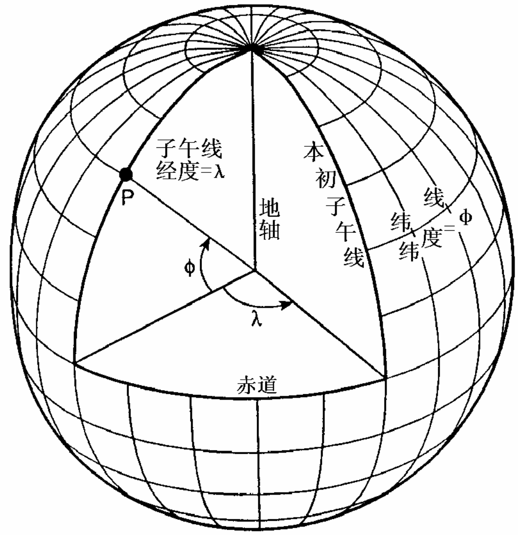


图 4-4：地球的经线和纬线

2．1．1 纬度（Latitude）

设椭球面上有一点 P（图 4-4），通过 P 点作椭球面的垂线，称之为过 P 点的法线。法线与赤道面的交角，叫做 P 点的地理纬度（简称纬度），通常以字母 ϕ表示。纬度从赤道起算，在赤道上纬度为 0 度，纬线离赤道愈远，纬度愈大，至极点纬度为 90 度。赤道以北叫北纬、以南叫南纬。

2．1．2 经度（Longitude）

过 P 点的子午面与通过英国格林尼治天文台的子午面所夹的二面角，叫做 P 点的地理经度（简称经度），通常用字母 λ 表示。国际规定通过英国格林尼治天文台的子午线为本初子午线（或叫首子午线），作为计算经度的起点，该线的经度为 0 度，向东 0-180 度叫东经，向西 0-180 度叫西经。

2．1．3 地面上点位的确定

地面上任一点的位置，通常用经度和纬度来决定。经线和纬线是地球表面上两组正交（相交为 90 度）的曲线，这两组正交的曲线构成的坐标，称为地理坐标系。地表面某两点经度值之差称为经差，某两点纬度值之差称为纬差。例如北京在地球上的位置可由北纬 39° 56’和东经 116° 24’来确定。

2．2 平面上的坐标系

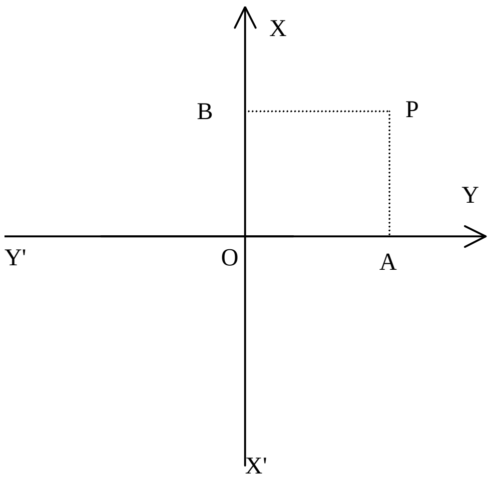
地理坐标是一种球面坐标。由于地球表面是不可展开的曲面，也就是说曲面上的各点不能直接表示在平面上，因此必须运用地图投影的方法，建立地球表面和平面上点的函数关系，使地球表面上任一点由地理坐标（ϕ、λ）确定的点，在平面上必有一个与它相对应的点，平面上任一点的位置可以用极坐标或直角坐标表示。

2．2．1 平面直角坐标系的建立

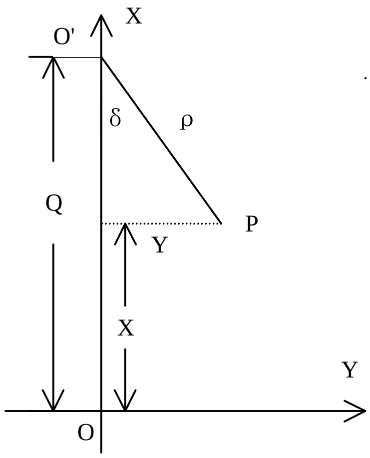
在平面上选一点 *O* 为直角坐标原点，过该点 *O* 作相互垂直的两轴 *X'OX* 和 *Y'OY* 而建立平面直角坐标系，如图 5 所示。

直角坐标系中，规定 *OX*、*OY* 方向为正值，*OX*、*OY* 方向为负值，因此在坐标系中的一个已知点 *P*，它的位置便可由该点对 *OX* 与 *OY* 轴的垂线长度唯一地确定，即 *x=AP*，*y=BP*，通常记为 *P(x，y)*。

2．2．2 平面极坐标系（Polar Coordinate）的建立



平面直角坐标系



平面极坐标系

图 4-5：平面直角坐标系和极坐标系

如图 5 所示，设 *O'* 为极坐标原点，*O'O* 为极轴，*P* 是坐标系中的一个点，则 *O'P* 称为极距，用符号 ρ 表示，即 $\rho=O'P$。$\angle OO'P$ 为极角，用符号 δ 表示，则 $\angle OO'P=\delta$。极角 δ 由极轴起算，按逆时针方向为正，顺时针方向为负。

极坐标与平面直角坐标之间可建立一定的关系式。由图 5 可知，直角坐标的 x 轴与极轴重合，二坐标系原点间距离 *OO'* 用 *Q* 表示，则有：

$X=Q-\rho\cos\delta$

$Y=\rho\sin\delta$

2．3 直角坐标系的平移和旋转

2．3．1 坐标系平移

如图 4-6 所示，坐标系 *XOY* 与坐标系 *X'O'Y'* 相应的坐标轴彼此平行，并且具有相同的正向。坐标系 *X'O'Y'* 是由坐标系 *XOY* 平行移动而得到的。设 *P* 点在坐标系 *XOY* 中的坐标为 *(x，y)*，在 *X'O'Y'* 中坐标为 *(x'，y')*，而 *(a，b)* 是 *O'* 在坐标系 *XOY* 中的坐标，于是：

$x=x'+a$

$y=y'+b$

上式即一点在坐标系平移前后之坐标关系式。

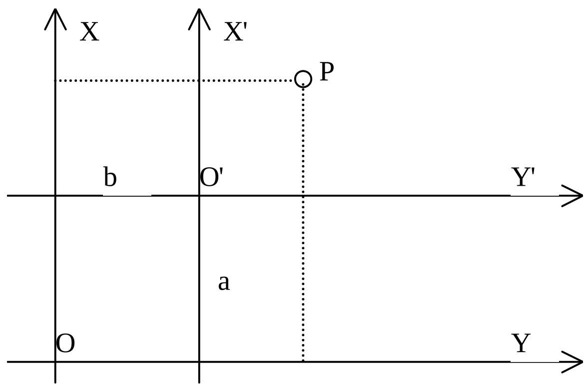


图 4-6：坐标平移

2．3．2 坐标系旋转

如图 4-7 所示，如坐标系 *XOY* 与坐标系 *X'O'Y'* 的原点重合，且对应的两坐标轴夹角为 θ，坐标系 *X'O'Y'* 是由坐标系 *XOY* 以 O 为中心逆时针旋转 θ 角后得到的。

$x=x'\cos\theta+y'\sin\theta$

$y=y'\cos\theta-x'\sin\theta$

上式即为经过旋转θ角后的二直角坐标系中某一点坐标的关系式。

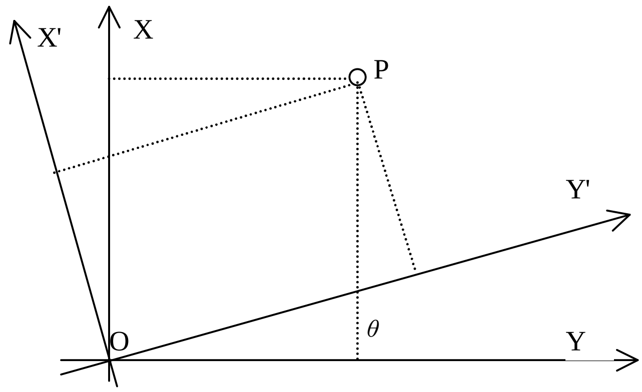


图 4-7：坐标旋转

2．3．3 坐标系平移和旋转

如图 4-8 所示，坐标系 *X'O'Y'* 的原点在坐标系 *XOY* 中的坐标为 a、b，*X* 轴与 *X'* 轴之夹角为θ。可以认为坐标系 *X'O'Y'* 原是与坐标系 XOY 重合，后因为 O' 分别平移了 a、b 之距离，并且坐标系二坐标轴 *O'X'* 与 *O'Y'* 又相对 *OX* 与 *OY* 逆时针旋转了 θ 角而得到的。

在二坐标系之间引入一个辅助坐标系 *X"O'Y"*，使它的二坐标轴 *O'X"* 与 *O'Y"* 分别与 *OX*、*OY* 平行。

在 *X"O'Y"* 系中有一点 P，其坐标为 (x'', y'')，则由坐标系平移公式与坐标系旋转公式可得：

$x=x''+a$

$y=y''+b$

故有

$x''=x'\cos\theta+y'\sin\theta$

$y''=y'\cos\theta-x'\sin\theta$

即

$x=x'\cos\theta+y'\sin\theta+a$

$y''=y'\cos\theta-x'\sin\theta+b$

上式即坐标系平移和旋转后新、旧坐标系中某一点坐标之关系式。

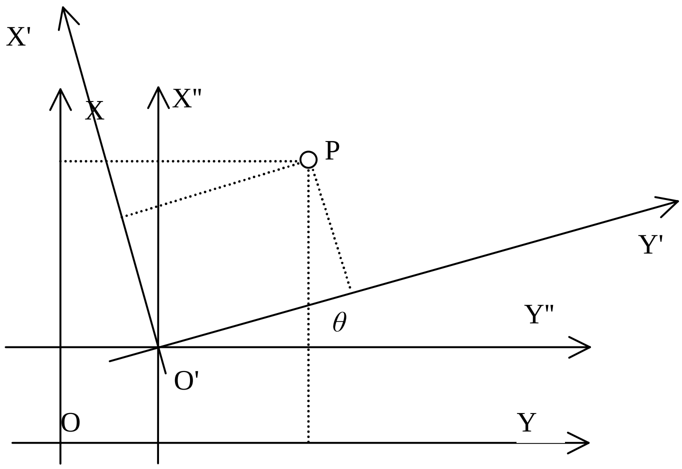


图 4-8：坐标平移和旋转

3．地图投影的基本问题

3．1 地图投影的概念

在数学中，投影（Project）的含义是指建立两个点集间一一对应的映射关系。同样，在地图学中，地图投影就是指建立地球表面上的点与投影平面上点之间的一一对应关系。地图投影的基本问题就是利用一定的数学法则把地球表面上的经纬线网表示到平面上。凡是地理信息系统就必然要考虑到地图投影，地图投影的使用保证了空间信息在地域上的联系和完整性，在各类地理信息系统的建立过程中，选择适当的地图投影系统是首先要考虑的问题。由于地球椭球体表面是曲面，而地图通常是要绘制在平面图纸上，因此制图时首先要把曲面展为平面，然而球面是个不可展的曲面，即把它直接展为平面时，不可能不发生破裂或褶皱。若用这种具有破裂或褶皱的平面绘制地图，显然是不实际的，所以必须采用特殊的方法将曲面展开，使其成为没有破裂或褶皱的平面。

3．2 地图投影的变形

3．2．1 变形的种类

地图投影的方法很多，用不同的投影方法得到的经纬线网形式不同。用地图投影的方法将球面展为平面，虽然可以保持图形的完整和连续，但它们与球面上的经纬线网形状并不完全相似。这表明投影之后，地图上的经纬线网发生了变形，因而根据地理坐标展绘在地图上的各种地面事物，也必然随之发生变形。这种变形使地面事物的几何特性（长度、方向、面积）受到破坏。把地图上的经纬线网与地球仪上的经纬线网进行比较，可以发现变形表现在长度、面积和角度三个方面，分别用长度比、面积比的变化显示投影中长度变形和面积变形。如果长度变形或面积变形为零，则没有长度变形或没有面积变形。角度变形即某一角度投影后角值与它在地球表面上固有角值之差。

1）长度变形

即地图上的经纬线长度与地球仪上的经纬线长度特点并不完全相同，地图上的经纬线长度并非都是按照同一比例缩小的，这表明地图上具有长度变形。

在地球仪上经纬线的长度具有下列特点：第一，纬线长度不等，其中赤道最长，纬度越高，纬线越短，极地的纬线长度为零；第二，在同一条纬线上，经差相同的纬线弧长相等；第三，所有的经线长度都相等。长度变形的情况因投影而异。在同一投影上，长度变形不仅随地点而改变，在同一点上还因方向不同而不同。

2）面积变形

即由于地图上经纬线网格面积与地球仪经纬线网格面积的特点不同，在地图上经纬线网格面积不是按照同一比例缩小的，这表明地图上具有面积变形。

在地球仪上经纬线网格的面积具有下列特点：第一，在同一纬度带内，经差相同的网络面积相等。第二，在同一经度带内，纬线越高，网络面积越小。然而地图上却并非完全如此。如在图 4-9-a 上，同一纬度带内，纬差相等的网格面积相等，这些面积不是按照同一比例缩小的。纬度越高，面积比例越大。在图 4-9-b 上，同一纬度带内，经差相同的网格面积不等，这表明面积比例随经度的变化而变化了。由于地图上经纬线网格面积与地球仪上经纬线网格面积的特点不同，在地图上经纬线网格面积不是按照同一比例缩小的，这表明地图上具有面积变形。面积变形的情况因投影而异。在同一投影上，面积变形因地点的不同而不同。

3）角度变形

是指地图上两条所夹的角度不等于球面上相应的角度，如在图 4-9-b 和图 4-9-c 上，只有中央经线和各纬线相交成直角，其余的经线和纬线均不呈直角相交，而在地球仪上经线和纬线处处都呈直角相交，这表明地图上有了角度变形。角度变形的情况因投影而异。在同一投影图上，角度变形因地点而变。

地图投影的变形随地点的改变而改变，因此在一幅地图上，就很难笼统地说它有什么变形，变形有多大。

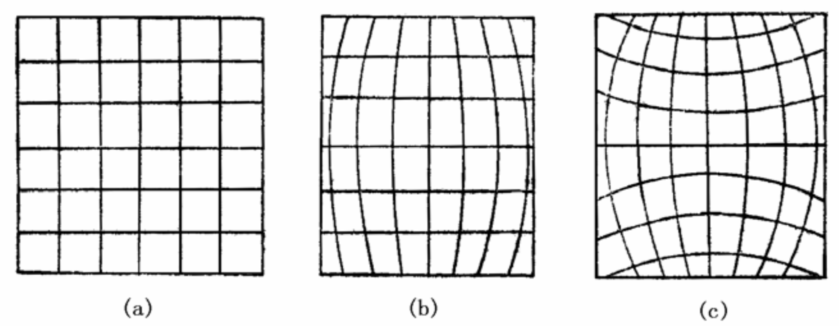


图 4-9：地图投影变形

3．2．2 变形椭圆

变形椭圆是显示变形的几何图形，从图 4-9 可以看到，实地上同样大小的经纬线在投影面上变成形状和大小都不相同的图形（比较图 4-9 中三个格网）。实际中每种投影的变形各不相同，通过考察地球表面上一个微小的圆形（称为微分圆）在投影中的表象——变形椭圆的形状和大小，就可以反映出投影中变形的差异（图 4-10）。

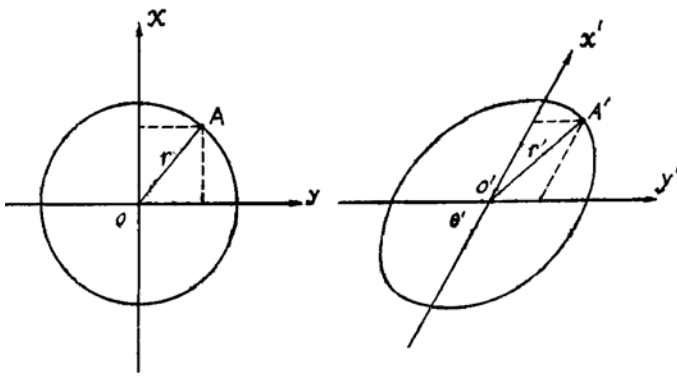


图 4-10：微分圆表示投影变形

3．3 地图投影的分类

地图投影的种类很多，为了学习和研究的方便，应对其进行分类。由于分类的标志不同，分类方法就不同。从使用地图的角度出发，需要了解下述几种分类。

3．3．1 按变形性质分类

按变形性质地图投影可以分为三类：等角投影、等积投影和任意投影。

1）等角投影

定义为任何点上二微分线段组成的角度投影前后保持不变，亦即投影前后对应的微分面积保持图形相似，故可称为正形投影。投影面上某点的任意两方向线夹角与椭球面上相应两线段夹角相等，即角度变形为零。等角投影在一点上任意方向的长度比都相等，但在不同地点长度比是不同的，即不同地点上的变形椭圆大小不同。

2）等积投影

定义为某一微分面积投影前后保持相等，亦即其面积比为 1，即在投影平面上任意一块

面积与椭球面上相应的面积相等，即面积变形等于零。

3）等距投影

在任意投影上，长度、面积和角度都有变形，它既不等角又不等积。但是在任意投影中，有一种比较常见的等距投影，定义为沿某一特定方向的距离，投影前后保持不变，即沿着该特定方向长度比为 1。在这种投影图上并不是不存在长度变形，它只是在特定方向上没有长度变形。等距投影的面积变形小于等角投影，角度变形小于等积投影。任意投影多用于要求面积变形不大、角度变形也不大的地图，如一般参考用图和教学地图。经过投影后地图上所产生的长度变形、面积变形和角度变形，是相互联系相互影响的。它们之间的关系是：在等积投影上不能保持等角特性，在等角投影上不能保持等积特性；在任意投影上不能保持等角和等积的特性；等积投影的形状变形比较大，等角投影的面积变形比较大。

3．3．2 按构成方法分类

地图投影最初建立在透视的几何原理上，它是把椭球面直接透视到平面上，或透视到可展开的曲面上，如圆柱面和圆锥面。圆柱面和圆锥面虽然不是平面，但可以展为平面。这样就得到具有几何意义的方位、圆柱和圆锥投影。随着科学的发展，为了使地图上变形尽量减小，或者为了使地图满足某些特定要求，地图投影就逐渐跳出了原来借助于几何面构成投影的框子，而产生了一系列按照数学条件构成的投影。因此，按照构成方法，可以把地图投影分为两大类：几何投影和非几何投影。

1）几何投影

几何投影是把椭球面上的经纬线网投影到几何面上，然后将几何面展为平面而得到。根据几何面的形状，可以进一步分为下述几类（图 4-11）：

（1．1）**方位投影**：以平面作为投影面，使平面与球面相切或相割，将球面上的经纬线投影到平面上而成。

（1．2）**圆柱投影**：以圆柱面作为投影面，使圆柱面与球面相切或相割，将球面上的经纬线投影到圆柱面上，然后将圆柱面展为平面而成。

（1．3）**圆锥投影**：以圆锥面作为投影面，使圆锥面与球面相切或相割，将球面上的经纬线投影到圆锥面上，然后将圆锥面展为平面而成。这里，我们可将方位投影看作圆锥投影的一种特殊情况，假设当圆锥顶角扩大到 180 度时，这圆锥面就成为一个平面，再将地球椭球体上的经纬线投影到此平面上。圆柱投影，从几何定义上讲，也是圆锥投影的一个特殊情况，设想圆锥顶点延伸到无穷远时，即成为一个圆柱。

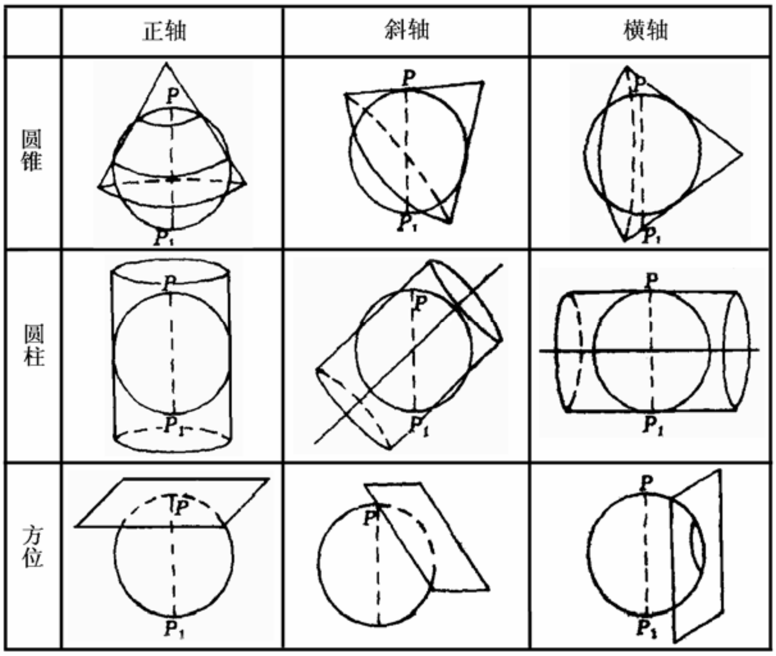


图 4-11：各种几何投影

2）非几何投影

不借助几何面，根据某些条件用数学解析法确定球面与平面之间点与点的函数关系。在这类投影中，一般按经纬线形状又分为下述几类：

（2．1）**伪方位投影**：纬线为同心圆，中央经线为直线，其余的经线均为对称于中央经线的曲线，且相交于纬线的共同圆心。

（2．2）**伪圆柱投影**：纬线为平行直线，中央经线为直线，其余的经线均为对称于中央经线的曲线。

（2．3）**伪圆锥投影**：纬线为同心圆弧，中央经线为直线，其余经线均为对称于中央经线的曲线。

（2．4）**多圆锥投影**：纬线为同周圆弧，其圆心均为于中央经线上，中央经线为直线，其余的经线均为对称于中央经线的曲线。

3．3．3 按照投影面积与地球相割或相切分类

1）割投影

以平面、圆柱面或圆锥面作为投影面，使投影面与球面相割，将球面上的经纬线投影到平面上、圆柱面上或圆锥面上，然后将该投影面展为平面而成。

2）切投影

以平面、圆柱面或圆锥面作为投影面，使投影面与球面相切，将球面上的经纬线投影到平面上、圆柱面上或圆锥面上，然后将该投影面展为平面而成。

3．4 地图投影的选择

地图投影选择得是否恰当，直接影响地图的精度和使用价值。这里所讲的地图投影选择，主要指中、小比例尺地图，不包括国家基本比例尺地形图。因为国家基本比例尺地形图的投影、分幅等，是由国家测绘主管部门研究制订*，不容许任意改变的，另外编制小区域大比例尺地图，无论采用什么投影，变形都是很小的。

选择制图投影时，主要要考虑以下因素：制图区域的范围、形状和地理位置，地图的用途、出版方式及其他特殊要求等，其中制图区域的范围、形状和地理位置是主要因素。

对于世界地图，常用的主要是正圆柱、伪圆柱和多圆锥投影。在世界地图中常用墨卡托投影绘制世界航线图、世界交通图与世界时区图；

我国出版的世界地图多采用等差分纬线多圆锥投影，选用这个投影，对于表现中国形状以及与四邻的对比关系较好，但投影的边缘地区变形较大。

对于半球地图，东、西半球图常选用横轴方位投影；南、北半球图常选用正轴方位投影；水、陆半球图一般选用斜轴方位投影。

对于其他的中、小范围的投影选择，须考虑到它的轮廓形状和地理位置，最好是使等变形线与制图区域的轮廓形状基本一致，以便减少图上变形。因此，圆形地区一般适于采用方位投影，在两极附近则采用正轴方位投影，以赤道为中心的地区采用横轴方位投影，在中纬度地区采用斜轴方位投影。在东西延伸的中纬度地区，一般多采用正轴圆锥投影，如中国与美国。在赤道两侧东西延伸的地区，则宜采用正轴圆柱投影，如印度尼西亚。在南北方向延伸的地区，一般采用横轴圆柱投影和多圆锥投影，如智利与阿根廷。

3．5 常用的一些地图投影

3．5．1 世界地图的投影

世界地图的投影主要考虑要保证全球整体变形不大，根据不同的要求，需要具有等角或等积性质，主要包括：等差分纬线多圆锥投影、正切差分纬线多圆锥投影（1976 年方案）、任意伪圆柱投影、正轴等角割圆柱投影。

3．5．2 半球地图的投影

东、西半球有横轴等面积方位投影、横轴等角方位投影；南、北半球有正轴等面积方位投影、正轴等角方位投影、正轴等距离方位投影。

3．5．3 各大洲地图投影

1）亚洲地图的投影：斜轴等面积方位投影、彭纳投影。

2）欧洲地图的投影：斜轴等面积方位投影、正轴等角圆锥投影。

3）北美洲地图的投影：斜轴等面积方位投影、彭纳投影。

---

* 目前我国采用的就是后面提及的高斯——克吕格投影及分幅方案。

4）南美洲地图的投影：斜轴等面积方位投影、桑逊投影。

5）澳洲地图的投影：斜轴等面积方位投影、正轴等角圆锥投影。

6）拉丁美洲地图的投影：斜轴等面积方位投影。

3．5．4 中国各种地图投影

1）中国全国地图投影：斜轴等面积方位投影、斜轴等角方位投影、彭纳投影、伪方位投影、正轴等面积割圆锥投影、正轴等角割圆锥投影。

2）中国分省（区）地图的投影：正轴等角割圆锥投影、正轴等面积割圆锥投影、正轴等角圆柱投影、高斯-克吕格投影（宽带）。

3）中国大比例尺地图的投影：多面体投影（北洋军阀时期）、等角割圆锥投影（兰勃特投影）（解放前）、高斯-克吕格投影（解放以后）。

4．高斯——克吕格投影

由于这个投影是由德国数学家、物理学家、天文学家高斯于 19 世纪 20 年代拟定，后经德国大地测量学家克吕格于 1912 年对投影公式加以补充，故称为高斯——克吕格投影。

高斯——克吕格投影在英美国家称为横轴墨卡托投影。美国编制世界各地军用地图和地球资源卫星象片所采用的全球横轴墨卡托投影（UTM）是横轴墨卡托投影的一种变型。高斯克吕格投影的中央经线长度比等于 1，UTM 投影规定中央经线长度比为 0.9996。在 6 度带内最大长度变形不超过 0.04%。

高斯克吕格投影的中央经线和赤道为互相垂直的直线，其他经线均为凹向并对称于中央经线的曲线，其他纬线均为以赤道为对称轴的向两极弯曲的曲线，经纬线成直角相交。在这个投影上，角度没有变形。中央经线长度比等于 1，没有长度变形，其余经线长度比均大于 1，长度变形为正，距中央经线愈远变形愈大，最大变形在边缘经线与赤道的交点上；面积变形也是距中央经线愈远，变形愈大。为了保证地图的精度，采用分带投影方法，即将投影范围的东西界加以限制，使其变形不超过一定的限度，这样把许多带结合起来，可成为整个区域的投影（图 4-12）。高斯——克吕格投影的变形特征是：在同一条经线上，长度变形随纬度的降低而增大，在赤道处为最大；在同一条纬线上，长度变形随经差的增加而增大，且增大速度较快。在 6 度带范围内，长度最大变形不超过 0.14%。

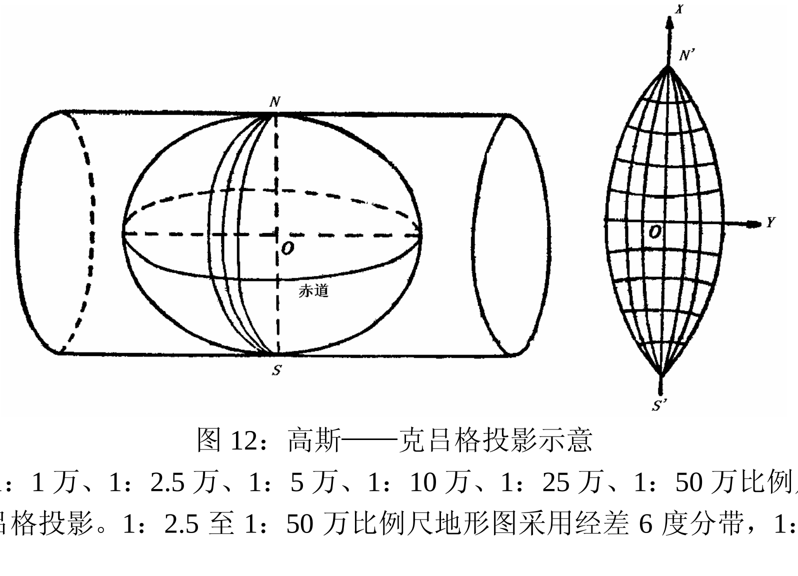


图 12：高斯——克吕格投影示意

我国规定 1：1 万、1：2.5 万、1：5 万、1：10 万、1：25 万、1：50 万比例尺地形图，均采用高斯克吕格投影。1：2.5 至 1：50 万比例尺地形图采用经差 6 度分带，1：1 万比例

尺地形图采用经差 3 度分带。

6 度带是从 0 度子午线起，自西向东每隔经差 6 为一投影带，全球分为 60 带，各带的带号用自然序数 1，2，3，…60 表示。即以东经 0-6 为第 1 带，其中央经线为 3E，东经 6-12 为第 2 带，其中央经线为 9E，其余类推（图 4-13）。

3 度带，是从东经 1 度 30 分的经线开始，每隔 3 度为一带，全球划分为 120 个投影带。图 4-13 表示出 6 度带与 3 度带的中央经线与带号的关系。

在高斯克吕格投影上，规定以中央经线为 X 轴，赤道为 Y 轴，两轴的交点为坐标原点。

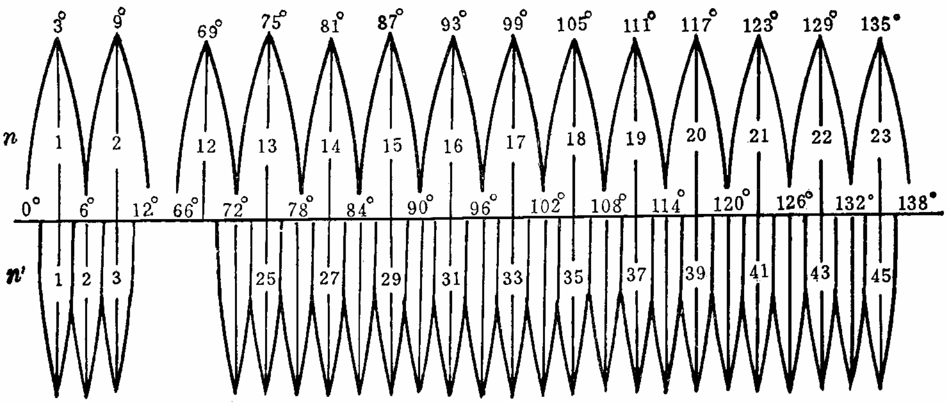


图 13：高斯——克吕格投影的分带

X 坐标值在赤道以北为正，以南为负；Y 坐标值在中央经线以东为正，以西为负。我国在北半球，X 坐标皆为正值。Y 坐标在中央经线以西为负值，运用起来很不方便。为了避免 Y 坐标出现负值，将各带的坐标纵轴西移 500 公里，即将所有 Y 值都加 500 公里。

由于采用了分带方法，各带的投影完全相同，某一坐标值（x，y），在每一投影带中均有一个，在全球则有 60 个同样的坐标值，不能确切表示该点的位置。因此，在 Y 值前，需冠以带号，这样的坐标称为通用坐标。

高斯克吕格投影各带是按相同经差划分的，只要计算出一带各点的坐标，其余各带都是适用的。这个投影的坐标值由国家测绘部门根据地形图比例尺系列，事先计算制成坐标表，供作业单位使用。

5．地形图的分幅和编号

国家基本比例尺地形图有 1：1 万、1：2.5 万、1：5 万、1：10 万、1：20 万、1：50 万和 1：100 万七种。普通地图通常按比例尺分为大、中、小三种，一般以 1：10 万和更大比例尺的地图称为大比例尺地图；1：10 万至 1：100 万的称为中比例尺地图；小于 1：100 万的称为小比例尺地图。对于一个国家或世界范围来讲，测制成套的各种比例尺地形图时，分幅编号尤其必要。通常这是由国家主管部门制定统一的图幅分幅和编号系统。

5．1 地形图的分幅

目前，我国采用的地形图分幅方案，是以 1：100 万地形图为基准，按照相同的经差和纬差定义更大比例尺地形图的分幅。百万分之一地图在纬度 0°—60^o之间的图幅，图幅大小按经差 6°，纬差 4^o分幅；在 60°—76^o之间的图幅，其经差为 12°，纬差为 4°；在 76°—80^o之间图幅的经差为 24°，纬差为 4°，所以各幅百万分之一地图都是经差 6°，纬差 4^o分幅的。

每幅百万分之一内各级较大比例尺地形图的划分，按规定的相应经纬差进行，其中，1：

50 万、1：20 万、1：10 万三种比例尺地形图，以百万分之一地图为基础直接划分。一幅百万分之一地形图划分四幅 1：50 万地形图，每幅为经差 3º，纬差 2º；一幅百万分之一地图划分为 36 幅 1：20 万地形图，每幅为经差 1º，纬差 40’；一幅百万分之一地图划分 144 幅 1:10 万地形图，每幅为经差 30’，纬差 20’。

每幅大于 1：10 万比例尺的地形图，则以 1：10 万图为基础进行逐级划分，一幅 1：10 万地形图划分四幅 1：5 万地形图；一幅 1：5 万地形图划分为四幅 1：2.5 万地形图。在 1：10 万图的基础上划分为 64 幅 1：1 万地形图；一幅 1：1 万地形图又划分为 4 幅 1：5000 地形图，（见表 4-1）。

表 4-1：基本比例尺地形图的图幅大小及其图幅间的数量关系。

比例尺（万）	图幅大小		图幅间的数量关系					
	经度	纬度						
1：100	6 度	4 度	1					
1：50	3 度	2 度	4	1				
1：20	1 度	40 分	36	9	1			
1：10	30 分	20 分	144	36	4	1		
1：5	15 分	10 分	576	144	16	4	1	
1：2.5	7.5 分	5 分	2304	576	64	16	4	1
1:1	3 分 45 秒	2.5 分	9216	2304	256	64	16	4

5．2 分幅编号

地形图的编号是根据各种比例尺地形图的分幅，对每一幅地图给予一个固定的号码，这种号码不能重复出现，并要保持一定的系统性。

地形图编号的最基本的方法是采用行列法，即把每幅图所在一定范围内的行数和列数组成一个号码。

5．2．1 1：100 万地图的编号

该种地形图的编号为全球统一分幅编号。

列数：由赤道起向南北两极每隔纬差 4º为一列，直到南北 88º（南北纬 88º至南北两极地区，采用极方位投影单独成图），将南北半球各划分为 22 列，分别用拉丁字母A、B、C、D.....V 表示。

行数：从经度 180º起向东每隔 6º为一行，绕地球一周共有 60 行，分别以数字 1、2、3、4......60 表示。

由于南北两半球的经度相同，规定在南半球的图号前加一个 S，北半球的图号前不加任何符号。一般来讲，把列数的字母写在前，行数的数字写在后，中间用一条短线连接。例如北京所在的一幅百万分之一地图的编号为 J-50（如图 14 所示）。

由于地球的经线向两极收敛，随着纬度的增加，同是 6º的经差但其纬线弧长已逐渐缩小，因此规定在纬度 60º-76º间的图幅采用双幅合并(经差为 12º,纬差为 4º);在纬度 76º-88º间的图幅采用四幅合并（经差为 24º，纬差为 4º）。这些合并图幅的编号，列数不变，行数（无论包含两个或四个）并列写在其后。例如北纬 80º-84º，西经 48º-72º的一幅百万分之一的地图编号应为U-19、20、21、22（图 4-14）。

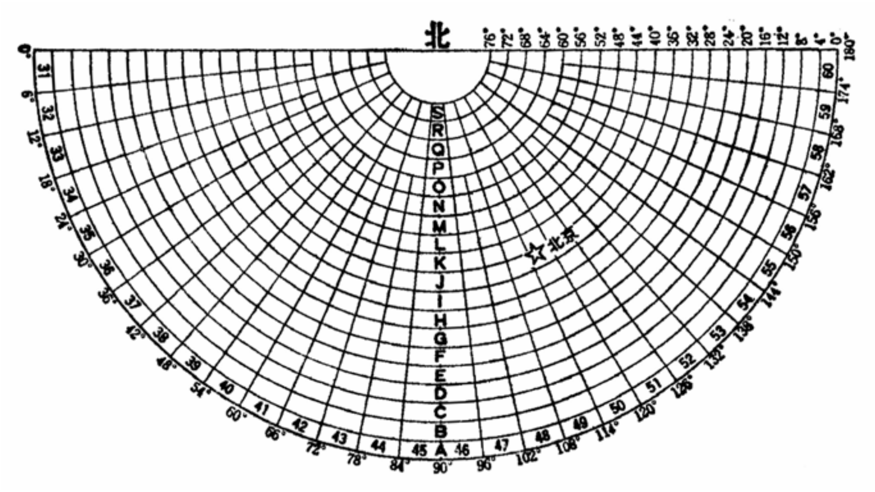


图 14：100 万地形图的分幅和编号（北半球）

5．2．2　1：50 万、1：20 万、1：10 万地形图的编号

一幅 1：100 万地图划分四幅 1：50 万地图，分别用甲、乙、丙、丁表示，其编号是在 1：100 万地形图的编号后加上它本身的序号，如 J-50-乙。

一幅 1：100 万地图划 36 幅 1：20 万地图，分别用带括号的数字（1）—（36）表示，其编号是在 1：100 万地形图的编号后加上它本身的序号，如 J-50-（28）。

一幅 1：100 万地图划分 144 幅 1：10 万地图，分别用数字 1—144 表示，其编号是在 1：100 万地形图的编号后加上它本身的序号，如 J-50-32。（图 4-15）

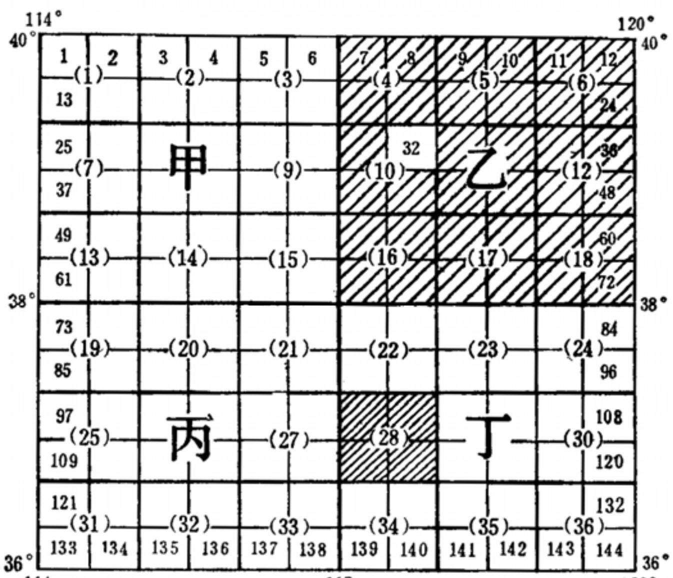


图 15：1：50 万、1：20 万、1：10 万地形图的分幅和编号示例

5．2．3　1：5 万、1：2.5 万、1：1 万地形图的编号

以 1：10 万地形图的编号为基础，将一幅 1：10 万地图划分四幅 1：5 万地图，分别用甲、乙、丙、丁表示，其编号是在 1：10 万地形图的编号后加上它本身的序号，如 J-50-32-

甲。再将一幅 1：5 万地图划分四幅 1：2.5 万地形图，分别用 1、2、3、4 表示，其编号是在 1：5 万地形图的编号后加上它本身的序号，如 J-50-32-甲-1。

1：1 万地形图的编号，是以一幅 1：10 万地形图划分为 64 幅 1：1 万地形图，分别以带括号的（1）—（64）表示，其编号是在 1：10 万图号后加上 1：1 万地图的序号，如 J-50-32-（10）。

一幅 1：1 万地形图划分为 4 幅 1：5000 地形图，分别用小写拉丁字母 a、b、c、d 表示，其编号是在 1：1 万图号后加上它本身的序号，如 J-50-32-（10）-a。

0(0),1(1),2(10),3(11),4(100),5(101),6(110),7(111),8(1000),9(1001)...

阿拉伯数字及其二级制表示

第五章 GIS 中的数据

导读：对空间数据的处理是 GIS 的核心功能，GIS 中的数据通常描述三部分信息：空间信息（位置，空间关系等等），非空间的属性信息，时间信息。数据按照其测量尺度，可以分为定名、定比、间隔、比率量四种类型。由于各方面的原因，GIS 数据存在着质量问题，了解数据质量有利于数据的正确使用，避免出现“Garbage In, Garbage Out”的情形。数据质量可以通过元数据进行描述，简单的说，元数据是对数据的描述，在 GIS 应用不断发展、空间数据不断增长的情况下，元数据有助于数据的共享和有效使用。

本章讲述了上述的内容。

1．数据涵义与数据类型

1．1 数据的涵义

1．1．1 数据

数据是用以载荷信息的载体。它可以是记录下来的某种可以识别的物理符号，数据的具体形式多种多样，如文本、图像、声音等都可以归入数据的范畴。虽然数据是信息的载体，但并非就是信息，只有理解了数据的含义、对数据做出解释，才能得到数据中所包含的信息。在计算机化的信息系统中，数据的格式往往与具体的计算机系统有关，随着载荷它的物理设备的形式而改变。信息系统对数据进行处理（运算、排序、编码、分类、增强等）就是为了得到数据中所包含的信息。

在地理信息系统中，由系统建立者输入、机器存储的各种专题地图和统计图表是数据；系统软件中所包含的代码是计算机系统中的二进制数据；用户对地理信息系统发出的各种指令也是数据，等等。因此，地理信息系统的建立和运行，就是信息或数据按一定的方式流动的过程。

数据的处理和解释是非常重要的环节。所谓数据处理，是指对数据进行收集、筛选、排序、归并、转换、检索、计算以及分析、模拟和预测的操作，其目的就是把数据转换成便于观察、分析、传输或进一步处理的形式；把数据加工成对正确管理和决策有用的数据；把数据编辑后存储起来，以供不断使用。数据处理是为了解释，而数据解释需要人的智慧、学识和经验。

1．1．2 地理数据（空间数据）的基本特征

地理数据一般具有三个基本特征：属性特征（非定位数据），表示实际现象或特征，例

如变量、级别、数量特征和名称等等。空间特征（定位数据）：表示现象的空间位置或现在所处的地理位置。空间特征又称为几何特征或定位特征，一般以坐标数据表示，例如笛卡尔坐标等。时间特征（时间尺度）：指现象或物体随时间的变化，其变化的周期有超短期的、短期的、中期的、长期的等等（图 5-1）。

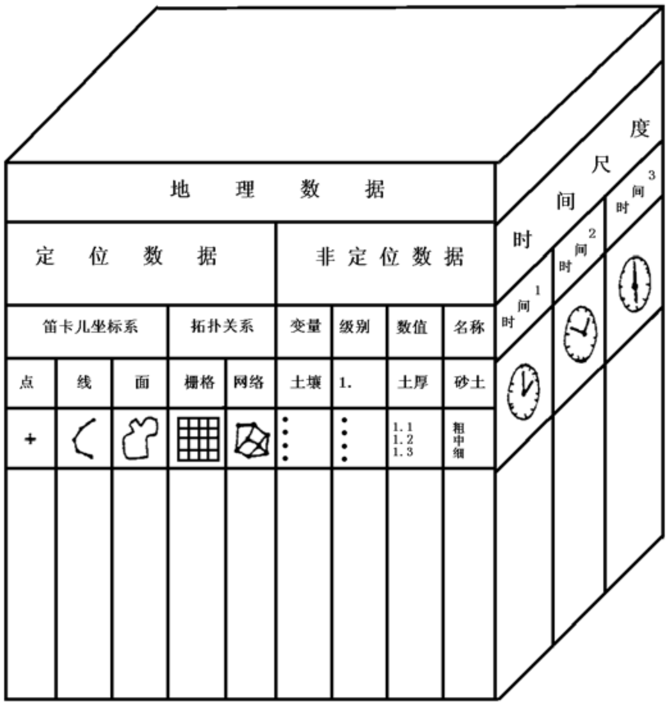


图 5-1：空间数据的基本特性[Jack Dangermond,1984]

1．2 空间数据的类型

在地理信息系统中，按照其特征，数据可分为三种类型：空间特征数据（定位数据）、时间属性数据（尺度数据）和专题属性数据（非定位数据）。对于绝大部分地理信息系统的应用来说，时间和专题属性数据结合在一起共同作为属性特征数据，而空间特征数据和属性特征数据统称为空间数据（或地理数据）。

1．2．1 空间特征数据

空间特征数据记录的是空间实体的位置、拓扑关系和几何特征，这是地理信息系统区别于其他数据库管理系统的标志。空间特征指空间物体的位置、形状和大小等几何特征，以及与相邻物体的拓扑关系。位置和拓扑特征是地理或空间信息系统所独有的，空间位置可以由不同的坐标系统来描述，如经纬度坐标、一些标准的地图投影坐标或是任意的直角坐标等。人类对空间目标的定位一般不是通过记忆其空间坐标，而是确定某一目标与其他更熟悉的目标间的空间位置关系，而这种关系往往也是拓扑关系。如一所学校位于哪个路口或哪条街道。

1．2．2 专题特征数据

专题特征指的是地理实体所具有的各种性质，如地形的坡度、坡向、某地的年降雨量、

土地酸碱类型、人口密度、交通流量、空气污染程度等。这类特征在其他类型的信息系统中均可存储和处理。专题属性特征通常以数字、符号、文本和图像等形式来表示。

1．2．3 时间特征数据

时间属性是指地理实体的时间变化或数据采集的时间等。严格地讲，空间数据总是在某一特定时间或时段内采集得到或计算产生的。由于有些空间数据随时间变化相对较慢，因而有时被忽略；有些时候，时间可以被看成一个专题特征。

1．3 空间数据的表示方法

一般地，表示地理现象的空间数据可以细分为：

类型数据：例如考古地点、道路线和土壤类型的分布等；

面域数据：例如随机多边形的中心点、行政区域界线和行政单元等；

网络数据：例如道路交点、街道和街区等；

样本数据：例如气象站、航线和野外样方的分布区等；

曲面数据：例如高程点、等高线和等值区域；

文本数据：例如地名、河流名称和区域名称；

符号数据：例如点状符号、线状符号和面状符号（晕线）等（如图 5-2 所示）。

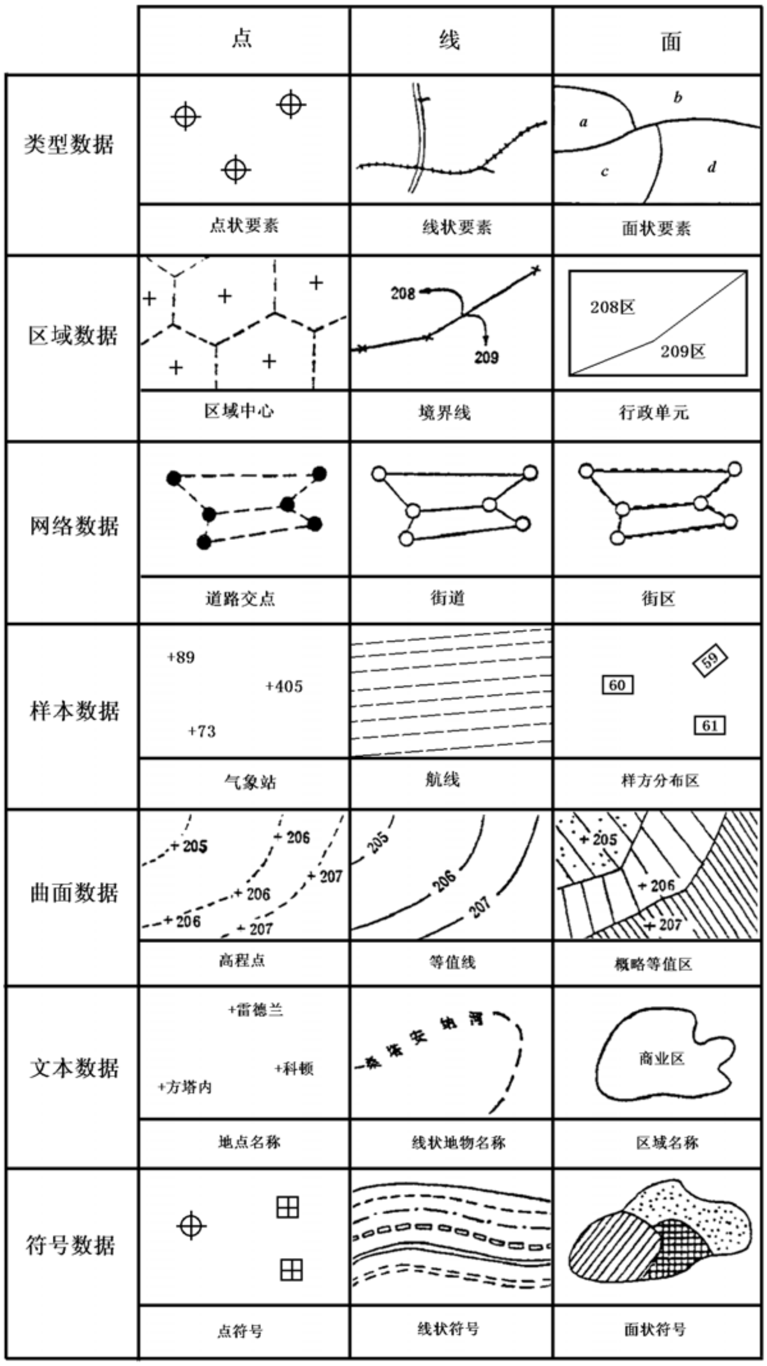


图 5-2：地理信息系统中各种数据以及其表现

2．数据的测量尺度

对特定现象的测量就是根据一定的标准对其赋值或打分。为了描述地理世界，对任何事物都要鉴别、分类和命名。这些都是量测的组成部分。它们所使用的参考标准或尺度是不同的。测量的尺度大致可以分成四个层次，由粗略至详细依次为：命名或类型、次序、间隔以及比例。

2．1 命名（Nominal）量

定性而非定量，不能进行任何算术运算，如一个城市的名字。命名式的测量尺度也称为类型测量尺度，只对特定现象进行标识，赋予一定的数值或符号而不定量描述。例如，可以用不同数值表示不同的土地利用类型、植被类型或岩石类型，但是这些数值之间无数量关系，对命名数据的逻辑运算只有“等于”或“不等于”两种形式，而其近似均值只能使用众数。

2．2 次序（Ordinal）量

线性坐标上不按值的大小，而是按顺序排列的数，例如，事故发生危险程度的级别由大到小被标为 1，2，3，…，级别的序号越低，其危险性越大，但危险性到底有多大并未给予定量的表达。序数值相互之间可以比较大小，但不能进行加、减、乘、除等算术运算。

次序测量尺度是基于对现象进行排序来标识的，如可以把山峰按高度分级为极高山、高山、中山、低山和丘陵等，将坡度分为陡、中、缓等。不同次序之间的间隔大小可以不同。对次序数据的逻辑运算除了“等于”与“不等于”之外，还可以比较它们的大小，即“大于”或“小于”。

2．3 间隔（Interval）量

不参照某个固定点，而是按间隔表示相对位置的数。按间隔量测的值相互之间可以比较大小，并且它们之间的差值大小是有意义的。

间隔测量尺度与比例测量尺度相似，但是间隔尺度的测量值无真的零值。例如，温度是间隔尺度的数据而不是比例数据，因为它的“0”测量值随着所使用的不同温度测量单位而不同。不能说 150F的温度是 75F的温度的两倍，因为这个比例在使用摄氏单位时就改变了。相反，降水量是比例数据，因为它有真的零值[*]。比例数据和间隔数据可用于加、减、乘、除等运算，而且可以求算术平均。

2．4 比率（Ratio）量

比例测量尺度的测量值指那些有真零值而且测量单位的间隔是相等的数据，比例测量尺度与使用的测量单位无关。

与某一固定点的比值计算，支持多种算术操作，如加、减、乘、除等。有关该类型属性域的例子很多，如年降雨量、海拔高度、人口密度、发病率等。

比例数据或间隔数据可以比较容易地被转变成次序或命名数据。而命名数据则很难被转化成次序、间隔数据或比例数据。由此可见，尽管命名数据或次序数据便于使用，易于理解，但有时不够精确，不能用于较高级的算术运算。而比例数据或间隔数据比较精确，便于计算机处理，但是在较复杂的 GIS 应用中，往往上述几种测量尺度的数据均需用到。

---

[*] 与比率量不同，间隔量中的零值往往是人为规定的，如摄氏温标中零摄氏度，时间中的零点等等。

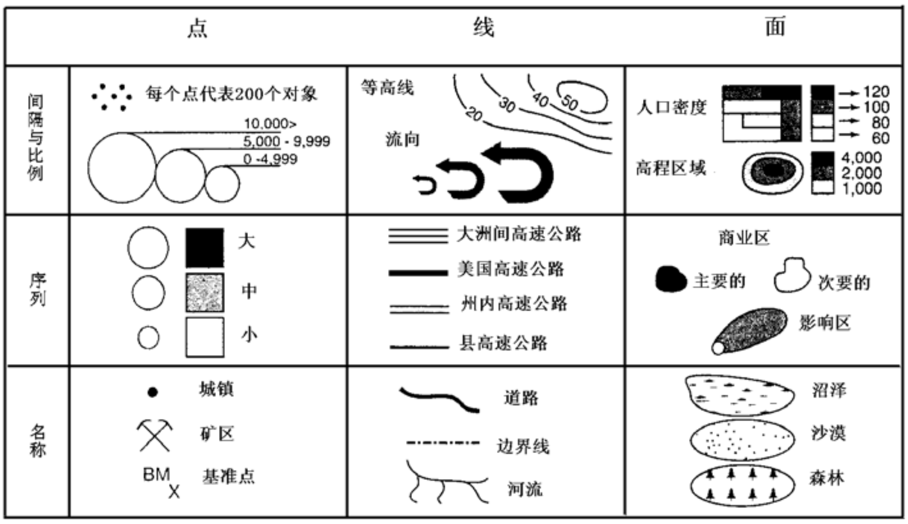


图 5-3：各种数据测量尺度以及其制图表现

3．地理信息系统的数据质量

地理信息系统是一个基于计算机软件、硬件和数据的集成系统，该系统主要通过空间及非空间数据的操作，实现空间检索、编辑及分析功能。在 GIS 的几个主要因素中，数据是一个极为重要的因素。在计算机软件、硬件环境选定之后，GIS 中数据质量的优劣，决定着系统分析质量以及整个应用的成败。GIS 提供的空间数据的分析方法被广泛用于各种领域，用于决策领域的数据，其质量要求应该是可知的或可预测的。

3．1 数据质量的基本概念

3．1．1 准确性（Accuracy）

即一个记录值（测量或观察值）与它的真实值之间的接近程度。这个概念是相当抽象的，似乎人们已经知道存在这样的事实。在实际中，测量的知识可能依赖于测量的类型和比例尺。一般而言，单个的观察或测量的准确性的估价仅仅是通过与可获得的最准确的测量或公认的分类进行比较。空间数据的准确性经常是根据所指的位置、拓扑或非空间属性来分类的。它可用误差（Error）来衡量。

3．1．2 精度（Precision）

即对现象描述的详细程度。如对同样的两点，精度低的数据并不一定准确度也低。精度要求测量能以最好的准确性来记录，但是这可能误导提供了较大的精度，因为超出一个测量仪器的已知准确度的数字在效率上是冗于的。因此，如果手工操作的数字化板所返回的坐标不可能依赖于比 0.1mm 还要准确的一个“真正的”数值，那么就不存在任何的点，在十分之一的地方是以 mm 表示的。

3．1．3 空间分辨率(Spatial Resolution)

分辨率是两个可测量数值之间最小的可辩识的差异。那么空间分辨率可以看作记录变化的最小距离。在一张用肉眼可读的地图上，假设一条线用来记录一个边界，分辨率通常由最小线的宽度来确定。地图上的线很少以小于 0.1mm 的宽度来画。在一个图形扫描仪中最细的物理分辨率从理论上讲是由设施的像元之间的分离来确定的。在一个激光打印机上这是一英寸的 300 分之一，而且在高质量的激光扫描仪上，这会细化十倍。如果没有放大，最细的激光扫描仪的线是看不到的，尽管这依赖于背景颜色的对照。因此，在人的视觉分辨率和设备物理分辨率之间存在着一个差异。一个相似的区别可以存在于两个最小距离之间，即当人操作者操作数字化仪时所区别的最小距离和数字化仪硬件可以不断地报告的最小距离。

3．1．4 比例尺（Scale）

比例尺是地图上一个记录的距离和它所表现的“真实世界的”距离之间的一个比例。地图的比例尺将决定地图上一条线的宽度所表现的地面的距离。例如，在一个 1：10000 比例尺的地图上，一条 0.5mm 宽度的线对应着 5m 的地面距离。如果这是线的最小的宽度，那么就不可能表示小于 5m 的现象。

3．1．5 误差（Error）

定义出一个所记录的测量和它的事实之间的准确性以后，很明显对于大多数目的而言，它的数值是不准确的。误差研究包括：位置误差，即点的位置的误差、线的位置的误差和多边形的位置的误差；属性误差；位置和属性误差之间的关系。

3．1．6 不确定性（Uncertainty）

地理信息系统的不确定性包括空间位置的不确定性、属性不确定性、时域不确定性、逻辑上的不一致性及数据的不完整性。空间位置的不确定性指 GIS 中某一被描述物体与其地面上真实物体位置上的差别；属性不确定性是指某一物体在 GIS 中被描述的属性与其真实的属性之差别；时域不确定性是指在描述地理现象时，时间描述上的差错；逻辑上的不一致性指数据结构内部的不一致性，尤其是指拓扑逻辑上的不一致性；数据的不完整性指对于给定的目标，GIS 没有尽可能完全地表达该物体。

3．2 空间数据质量问题的来源

从空间数据的形式表达到空间数据的生成，从空间数据的处理变换到空间数据的应用，在这两个过程中都会有数据质量问题的发生。下面按照空间数据自身存在的规律性，从几个方面来阐述空间数据质量问题的来源。

3．2．1 空间现象自身存在的不稳定性

空间数据质量问题首先来源于空间现象自身存在的不稳定性。空间现象自身存在的不稳

定性包括空间特征和过程在空间、专题和时间内容上的不确定性。空间现象在空间上的不确定性指其在空间位置分布上的不确定性变化；空间现象在时间上的不确定性表现为其在发生时间段上的游移性；空间现象在属性上的不确定性表现为属性类型划分的多样性，非数值型属性值表达的不精确性。因此，空间数据存在质量问题是不可避免的。

3．2．2 空间现象的表达

数据采集中的测量方法以及量测精度的选择等受到人类自身的认识和表达的影响，这对于数据的生成会出现误差。如在地图投影中，由椭球体到平面的投影转换必然产生误差；用于获取各种原始数据的各种测量仪器都有一定的设计精度，如 GPS 提供的地理位置数据都有用户要求的一定设计精度，因而数据误差的产生不可避免。

3．2．3 空间数据处理中的误差

在空间数据处理过程中，容易产生的误差有以下几种：

投影变换：地图投影是开口的三维地球椭球面到二维场平面的拓扑变换。在不同投影形式下，地理特征的位置、面积和方向的表现会有差异。

地图数字化和扫描后的矢量化处理：数字化过程采点的位置精度、空间分辨率、属性赋值等都可能出现误差。

数据格式转换：在矢量格式和栅格格式之间的数据格式转换中，数据所表达的空间特征的位置具有差异性。

数据抽象：在数据发生比例尺变换时，对数据进行的聚类、归并、合并等操作时产生的误差，如知识性误差和数据所表达的空间特征位置的变化误差。

建立拓扑关系：拓扑过程中伴随有数据所表达的空间特征的位置坐标的变化。

与主控数据层的匹配：一个数据库中，常存储同一地区的多层数据面，为保证各数据层之间空间位置的协调性，一般建立一个主控数据层以控制其它数据层的边界和控制点。在与主控数据层匹配的过程中也会存在空间位移，导致误差。

数据叠加操作和更新：数据在进行叠加运算以及数据更新时，会产生空间位置和属性值的差异。

数据集成处理：指在来源不同、类型不同的各种数据集的相互操作过程中所产生的误差。数据集成是包括数据预处理、数据集之间的相互运算、数据表达等过程在内的复杂过程，其中位置误差、属性误差都会出现。

数据的可视化表达：数据在可视化表达过程中为适应视觉效果，需对数据的空间特征位置、注记等进行调整，由此产生数据表达上的误差。

数据处理过程中误差的传递和扩散：在数据处理的各个过程中，误差是累计和扩散的，前一过程的累计误差可能成为下一个阶段的误差起源，从而导致新的误差的产生。

3．2．4 空间数据使用中的误差

在空间数据使用的过程中也会导致误差的出现，主要包括两个方面：一是对数据的解释过程，二是缺少文档。对于同一种空间数据来说，不同用户对它的内容的解释和理解可能不同，处理这类问题的方法是随空间数据提供各种相关的文档说明，如元数据。另外，缺少对某一地区不同来源的空间数据的说明，如缺少投影类型、数据定义等描述信息，这样往往导

致数据用户对数据的随意性使用而使误差扩散。

表 5-1：数据的主要误差来源

数据处理过程	误差来源
数据搜集	野外测量误差：仪器误差、记录误差 遥感数据误差：辐射和几何纠正误差、信息提取误差 地图数据误差：原始数据误差、坐标转换、制图综合及印刷
数据输入	数字化误差：仪器误差、操作误差 不同系统格式转换误差：栅格-矢量转换、三角网-等值线转换
数据存储	数值精度不够 空间精度不够：每个格网点太大、地图最小制图单元太大
数据处理	分类间隔不合理 多层数据叠合引起的误差传播：插值误差、多源数据综合分析误差 比例尺太小引起的误差
数据输出	输出设备不精确引起的误差 输出的媒介不稳定造成的误差
数据使用	对数据所包含的信息的误解 对数据信息使用不当

3．3 常见空间数据的误差分析

GIS 中的误差是指 GIS 中数据表示与其现实世界本身的差别。数据误差的类型可以是随机的，也可以是系统的。归纳起来，数据的误差主要有四大类，即几何误差、属性误差、时间误差和逻辑误差。在这几种误差中，属性误差和时间误差与普通信息系统中的误差概念是一致的，几何误差是地理信息系统所特有的，而几何误差、属性误差和时间误差都会造成逻辑误差，因此下面主要讨论逻辑误差和几何误差。

3．3．1 误差的类型

1）逻辑误差

数据的不完整性是通过上述四类误差反映出来的。事实上检查逻辑误差，有助于发现不完整的数据和其他三类误差。对数据进行质量控制或质量保证或质量评价，一般先从数据的逻辑性检查入手。如图 5-4 所示，其中桥或停车场等与道路是相接的，如果数据库中只有桥或停车场，而没有与道路相连，则说明道路数据被遗漏，使数据不完整。

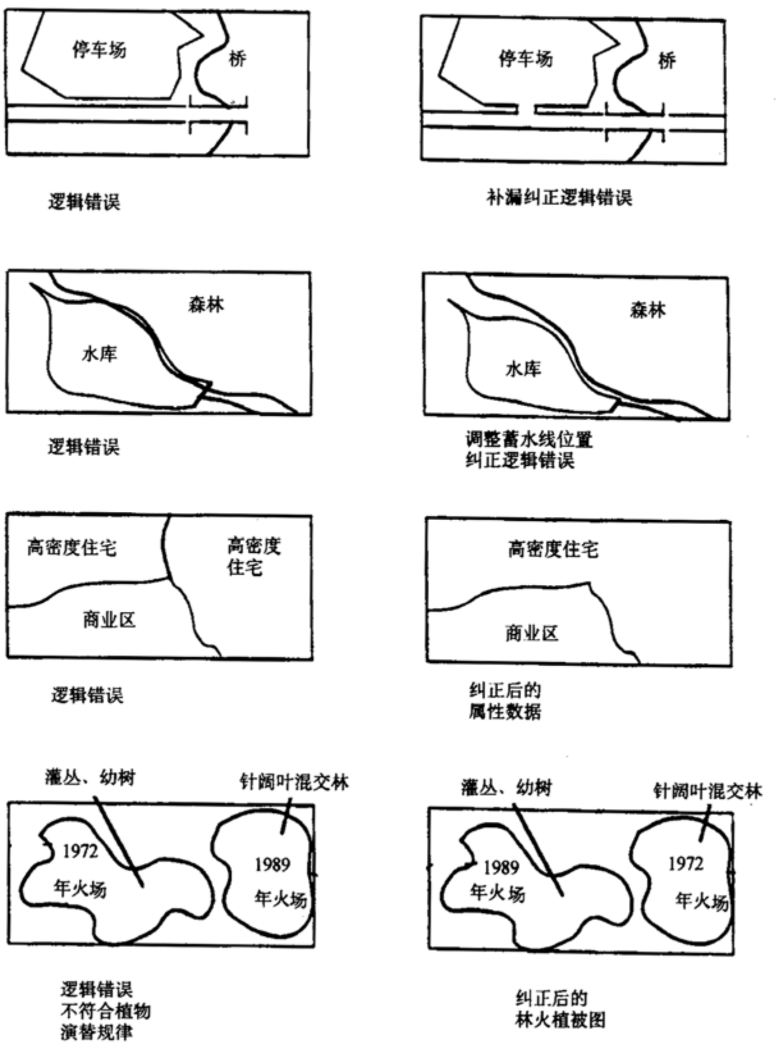


图 5-4：各种逻辑误差

2）几何误差

由于地图是以二维平面坐标表达位置，在二维平面上的几何误差主要反映在点和线上。

（2．1）点误差

关于某点的点误差即为测量位置（x，y）与其真实位置（x_0，y_0）的差异。真实位置的测量方法比测量位置的要更加精确，如在野外使用高精度的GPS方法得到。点误差可通过计算坐标误差和距离的方法得到。坐标误差定义为：

$\Delta x = x - x_0$

$\Delta y = y - y_0$

为了衡量整个数据采集区域或制图区域内的点误差，一般抽样测算（Δx, Δy）。抽样点应随机分布于数据采集区内，并具有代表性。这样抽样点越多，所测的误差分布就越接近于点误差的真实分布。

（2．2）线误差

线在地理信息系统数据库中既可表示线性现象，又可以通过连成的多边形表示面状现象。第一类是线上的点在真实世界中是可以找到的，如道路、河流、行政界线等，这类的线性特征的误差主要产生于测量和对数据的后处理；第二类是现实世界中找不到的，如按数学投影定义的经纬线、按高程绘制的等高线，或者是气候区划线和土壤类型界限等，这类线性特征的线误差及在确定线的界限时的误差，被称为解译误差。解译误差与属性误差直接相关，

若没有属性误差，则可以认为那些类型界线是准确的，因而解译误差为零。

另外，线分为直线、折线、曲线与直线混合的线（图 5-5）。GIS 数据库中用两种方法表达曲线、折线，图 5-6 对这两类误差作了对照。

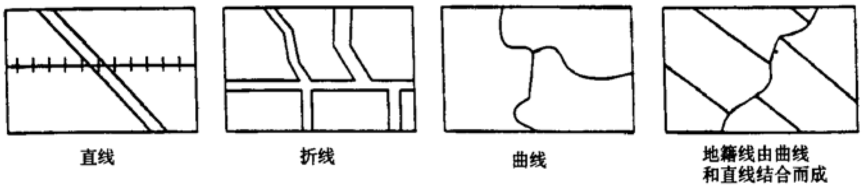


图 5-5：各种线（直线、折线、曲线）

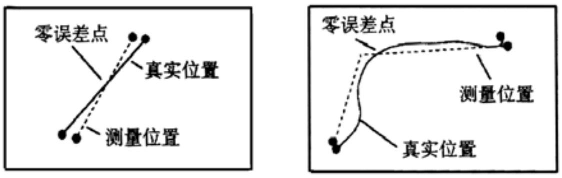


图 5-6：折线和曲线的误差

线误差分布可以用 Epsilon 带模型来描述，它由沿着一条线以及两侧定宽的带构成，真实的线以某一概率落于 Epsilon 带内。Epsilon 带是等宽的（类似于后面讲述的缓冲区，不过其意义不同），在此基础上，误差带模型被提出，与 Epsilon 带模型相比，它在中间最窄而在两端较宽。基于误差带模型，可以把直线与折线误差分布的特点分别看作是“骨头型”或者“车链型”的误差分布带模式（图 5-7）。

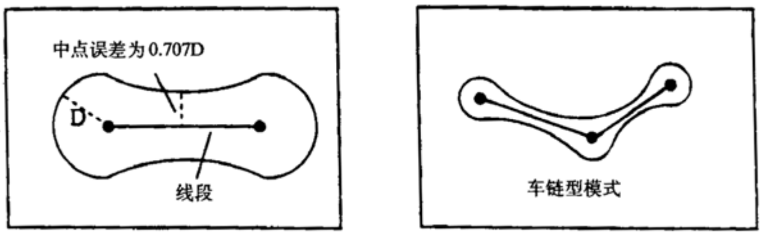


图 5-7：折线误差的分布

对于曲线的误差分布或许应当考虑“串肠型模式”（图 5-8）。

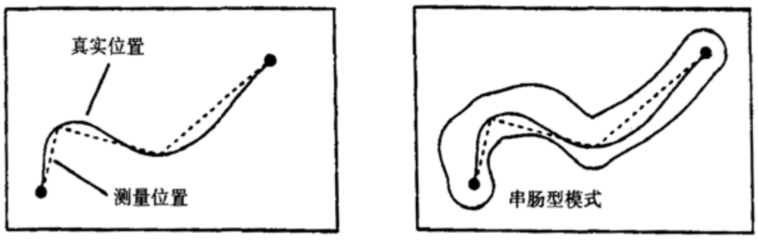


图 5-8：曲线的误差分布

3．3．2 地图数据的质量问题

地图数据是现有地图经过数字化或扫描处理后生成的数据。在地图数据质量问题中，不仅含有地图固有的误差，还包括图纸变形、图形数字化等误差。

1）地图固有误差：是指用于数字化的地图本身所带有的误差，包括控制点误差、投影误差等。由于这些误差间的关系很难确定，所以很难对其综合误差作出准确评价。如果假定综合误差与各类误差间存在线性关系，即可用误差传播定律来计算综合误差。

2）材料变形产生的误差：这类误差是由于图纸的大小受湿度和温度变化的影响而产生的。温度不变的情况下，若湿度由 0%增至 25%，则纸的尺寸可能改变 1.6%；纸的膨胀率和收缩率并不相同，即使湿度又恢复到原来的大小，图纸也不能恢复原有的尺寸，一张 6 英寸的图纸因湿度变化而产生的误差可能高达 0.576 英寸。在印刷过程中，纸张先随温度的升高而变长变宽，又由于冷却而产生收缩。

3）图象数字化误差：数字化方式主要有跟踪数字化和扫描数字化两种。跟踪数字化一般有点方式和流方式两种工作方式，前者在实际工作中使用较多，后者进行数字化所产生的误差要比前者大得多。

不同数据录入方式对数据质量的影响

跟踪数字化：

影响其数据质量的因素主要有：数字化要素对象、数字化操作人员、数字化仪和数字化操作。其中，数字化要素对象：地理要素图形本身的高度、密度和复杂程度对数字化结果的质量有着显著影响，如粗线比细线更易引起误差，复杂曲线比平直线更易引起误差，密集的要素比稀疏的要素更易引起误差等；数字化操作人员：数字化操作人员的技术与经验不同，所引入的数字化误差也会有较大的误差，这主要表现在最佳采点点位的选择、十字丝与目标重叠程度的判断能力等方面，另外，数字化操作人员的疲劳程度和数字化的速度也会影响数字化的质量；数字化仪的分辨率和精度对数字化的质量有着决定性的影响；数字化操作方式也会影响到数字化数据的质量，如曲线采点方式（流方式或点方式）和采点密度等。

扫描数字化：

扫描数字化采用高精度扫描仪将图形、图象等扫描并形成栅格数据文件，再利用扫描矢量化软件对栅格数据文件进行处理，将它转换为矢量图形数据。矢量化过程有两种方式：即交互式和全自动。影响扫描数字化数据质量的因素包括原图质量（如清晰度）、扫描精度、扫描分辨率、配准精度、校正精度等。

3．3．3 遥感数据的质量问题

遥感数据的质量问题，一部分来自遥感仪器的观测过程，一部分来自遥感图象处理和解译过程。遥感观测过程本身存在着精确度和准确度的限制，这一过程产生的误差主要表现为空间分辨率、几何畸变和辐射误差，这些误差将影响遥感数据的位置和属性精度。遥感图像处理和解译过程，主要产生空间位置和属性方面的误差。这是由图像处理中的影像或图像校正和匹配以及遥感解译判读和分类引入的，其中包括混合像元的解译判读所带来的属性误差。

3．3．4 测量数据的质量问题

测量数据主要指使用大地测量、GPS、城市测量、摄影测量和其他一些测量方法直接量测所得到的测量对象的空间位置信息。这部分数据质量问题，主要是空间数据的位置误差。空间数据的位置通常以坐标表示，空间数据位置的坐标与其经纬度表示之间存在着某误差因素，由于这种误差因素无法排除，一般也不作为误差考虑。测量方面的误差通常考虑的是系统误差、操作误差和偶然误差。

系统误差的发生与一个确定的系统有关，它受环境因素（如温度、湿度和气压等）、仪器结构与性能以及操作人员技能等方面的因素综合影响而产生。系统误差不能通过重复观测加以检查或消除，只能用数字模型模拟和估计。

操作误差是操作人员在使用设备、读书或记录观测值时，因粗心或操作不当而产生的。

应采用各种方法检查和消除操作误差。一般地，操作误差可通过简单的几何关系或代数检查验证其一致性，或通过重复观测检查并消除操作误差。

偶然误差是一种随机性的误差，由一些不可测和不可控的因素引入。这种误差具有一定的特征，如正负误差出现频率相同、大误差少、小误差多等。偶然误差可采用随机模型进行估计和处理。

3．4 空间数据质量控制

数据质量控制是个复杂的过程，要控制数据质量应从数据质量产生和扩散的所有过程和环节入手，分别用一定的方法减少误差。空间数据质量控制常见的方法有：

3．4．1 传统的手工方法

质量控制的人工方法主要是将数字化数据与数据源进行比较，图形部分的检查包括目视方法、绘制到透明图上与原图叠加比较，属性部分的检查采用与原属性逐个对比或其他比较方法。

3．4．2 元数据方法

数据集的元数据中包含了大量的有关数据质量的信息，通过它可以检查数据质量，同时元数据也记录了数据处理过程中质量的变化，通过跟踪元数据可以了解数据质量的状况和变化。

3．4．3 地理相关法

用空间数据的地理特征要素自身的相关性来分析数据的质量。如从地表自然特征的空间分布着手分析，山区河流应位于微地形的最低点，因此，叠加河流和等高线两层数据时，如河流的位置不在等高线的外凸连线上，则说明两层数据中必有一层数据有质量问题，如不能确定哪层数据有问题时，可以通过将它们分别与其它质量可靠的数据层叠加来进一步分析。因此，可以建立一个有关地理特征要素相关关系的知识库，以备各空间数据层之间地理特征要素的相关分析之用。

4．空间数据的元数据

Metadata 可以译成元数据，是描述数据的数据。在地理空间数据中，元数据是说明数据内容、质量、状况和其他有关特征的背景信息。元数据并不是一个新的概念。实际上传统的图书馆卡片、出版图书的版权说明、磁盘的标签等都是元数据。纸质地图的元数据主要表现为地图类型、地图图例，包括图名、空间参照系和图廓坐标、地图内容说明、比例尺和精度、编制出版单位和日期或更新日期、销售信息等。在这种形式下，元数据是可读的，生产者和用户之间容易交流，用户通过它可以非常容易地确定该书或地图是否能够满足其应用的需要。

随着计算机技术和GIS技术发展，特别是网络通信技术的发展，空间数据共享日益普遍。

管理和访问大型数据集的复杂性正成为数据生产者和用户面临的突出问题。数据生产者需要有效的数据管理和维护办法；用户需要找到更快、更加全面和有效的方法，以便发现、访问、获取和使用现势性强、精度高、易管理和易访问的地理空间数据。在这种情况下，空间数据的内容、质量、状况等元数据信息变得更加重要，成为信息资源有效管理和应用的重要手段。地理信息元数据标准和操作工具已经成为国家空间数据基础设施*的一个重要组成部分。

在地理信息系统应用中，元数据的主要作用可以归纳为如下几个方面：

1）帮助数据生产单位有效地管理和维护空间数据、建立数据文档，并保证即使其主要工作人员离退时，也不会失去对数据情况的了解；

2）提供有关数据生产单位数据存储、数据分类、数据内容、数据质量、数据交换网络及数据销售等方面的信息，便于用户查询检索地理空间数据；

3）帮助用户了解数据，以便就数据是否能满足其需求做出正确的判断；

4）提供有关信息，以便用户处理和转换有用的数据。

可见，元数据是使数据充分发挥作用的重要条件之一，它可以用于许多方面，包括数据文档建立、数据发布、数据浏览、数据转换等。元数据对于促进数据的管理、使用和共享均有重要的作用。

4．1 元数据的概念及类型

4．1．1 元数据的概念

元数据是关于数据的描述性数据信息，它应尽可能多地反映数据集自身的特征规律，以便于用户对数据集的准确、高效与充分的开发与利用，不同领域的数据库，其元数据的内容会有很大差异。通过元数据可以检索、访问数据库，可以有效利用计算机的系统资源，可以对数据进行加工处理和二次开发等。

到目前为止，科学界关于元数据认识的共同点是：元数据的目的就是促进数据集的高效利用，并为计算机辅助软件工程（CASE）服务。元数据的内容包括：

1）对数据集的描述；对数据集中各数据项、数据来源、数据所有者及数据序代（数据生产历史）等的说明；

2）对数据质量的描述，如数据精度、数据的逻辑一致性、数据完整性、分辨率、元数据的比例尺等；

3）对数据处理信息的说明，如量纲的转换等；

4）对数据转换方法的描述；

5）对数据库的更新、集成等的说明。

4．1．2 元数据的类型

元数据的分类研究的目的在于充分了解和更好地使用元数据。分类的原则不同，元数据的分类体系和内容将会有很大的差异。

1）根据元数据的内容分类

由于不同性质、不同领域的数据所需要的元数据内容有差异，而且为不同应用目的而建

---

* 国家空间数据基础设施的具体描述见“地球信息科学与数字地球”一章。

设的数据库的元数据内容会有很大的差异，所以将元数据化分为三种类型：

（1．1）科研型元数据：其主要目标是帮助用户获取各种来源的数据及其相关信息，它不仅包括如数据源名称、作者、主体内容等传统的、图书管理式的元数据，还包含数据拓扑关系等。这类元数据的任务是帮助科研工作者高效获取所需数据。

（1．2）评估型元数据：主要服务于数据利用的评价，内容包括数据最初收集情况、收集数据所用的仪器、数据获取的方法和依据、数据处理过程和算法、数据质量控制、采样方法、数据精度、数据的可信度、数据潜在应用领域等。

（1．3）模型元数据：用于描述数据模型的元数据与描述数据的元数据在结构上大致相同，其内容包括模型名称、模型类型、建模过程、模型参数、边界条件、作者、引用模型描述、建模使用软件、模型输出等。

2）根据元数据描述对象分类

（1．1）数据层元数据：指描述数据集中每个数据的元数据，内容包括日期邮戳、位置戳、量纲、注释、误差标识、缩略标识、存在问题标识、数据处理过程等。

（1．2）属性元数据：是关于属性数据的元数据，内容包括为表达数据及其含义所建的数据字典、数据处理规则（协议），如采样说明、数据传输线路及代数编码等。

（1．3）实体元数据：是描述整个数据集的元数据，内容包括数据集区域采样原则、数据库的有效期、数据时间跨度等。

3）根据元数据在系统中的作用分类

（1．1）系统级别元数据：指用于实现文件系统特征或管理文件系统中数据的信息，如访问数据的时间、数据的大小、在存储级别中的当前位置、如何存储数据块以保证服务控制质量等。

（1．2）应用层元数据：指有助于用户查找、评估、访问和管理数据等与数据用户有关的信息，如文本文件内容的摘要信息、图形快照、描述与其它数据文件相关关系的信息。它往往用于高层次的数据管理，用户通过它可以快速获取合适的数据。

4）根据元数据的作用分类

（4．1）说明元数据：是为用户使用数据服务的元数据。它一般用自然语言表达，如源数据覆盖的空间范围、源数据图的投影方式及比例尺的大小、数据集说明文件等，这类元数据多为描述性信息，侧重于数据库的说明。

（4．2）控制元数据：是用于计算机操作流程控制的元数据，这类元数据由一定的关键词和特定的句法来实现。其内容包括数据存储和检索文件、检索中与目标匹配方法、目标的检索和显示、分析查询结果排列显示、根据用户要求修改数据库中原有的内部顺序、数据转换方法、空间数据和属性数据的集成、根据索引项把数据绘制成图、数据模型的建设和利用等。这类元数据主要是与数据库操作有关的方法。

空间数据元数据中所用到的概念：

空间数据(Geospatial Data)：用于确定具有自然特征或者人工建筑特征的地理实体的地理位置、属性及其边界的信息；

类型(Type)：在元数据标准中，数据类型指该数据能接收的值的类型；

对象(Object)：对地理实体的部分或整体的数字表达；

实体类型(Entity Type)：对于具有相似地理特征的地理实体集合的定义和描述；

点(Point)：用于位置确定的零维地理对象；

结点(Node)：拓扑连接两个或多个链或环的一维对象；

标识点(Label Point)：显示地图后图表时用于特征标识的参考点；

线(Line)：一维对象的一般术语；

线段(Line Segment)：两个点之间的直线段；

串(String)：由相互连接的一系列线段组成的没有分支线段的序列，它可与自身或与其它线相切；

弧(Arc)：由数学表达式确定的点集组成的弧状曲线；

链(Link)：两个结点之间的拓扑关联；

链环(Chain)：非相切线段或由结点区分的弧段构成的有方向无分支序列；

环(Ring)：封闭状不相切链环或弧段序列；

多边形(Ploygon)：在二维平面中由封闭弧段包围的区域；

外多边形(Universe Polygon)：数据覆盖区域内最外侧的多边形，其面积是其它所有多边形的面积之和；

内部区域(Interior Area)：不包括其边界的区域；

格网(Grid)：组成一规则或近似规则的棋盘状镶嵌表面的格网集合，或者组成一规则或近似规则的棋盘状镶嵌的点集合；

格网单元(Grid Cell)：表示格网最小分为要素的二维对象；

矢量(Vector)：有方向线的组合；

栅格(Raster)：同一格网或数字影像的一个或多个叠加层；

像元(Pixel)：二维图形要素，它是数学影像最小要素；

栅格对象(Raster Object)：一个或多个影象或格网，每个影象或格网表示一个数据层，各层之间相应的格网单元或像元一致且相互套准；

图形(Graph)：与预定义的限制规则一致的零维（如 node）、一维(link 或 chain)和二维(T 多边形)有拓扑相关的对象集；

数据层(Layer)：集成到一起的面域分布空间数据集，它用于表示一个主体中的实体，或者有一公共属性或属性值的空间对象的联合；

层(Stratum)：在有序系统中数据层、级别或梯度序列；

纬度(Latitude)：在中央经线上度量，以角度单位度量离开赤道的距离；

经度(Longitude)：经线面到格林尼治中央经线面的角度距离；

经圈(Meridian)：穿过地球两极的地球的大圆圈；

坐标(Ordinate)：在笛卡尔坐标系中沿平行于 X 轴和 Y 轴测量的坐标值；

投影(Projection)：将地球球面坐标中的空间特征（集）转化到平面坐标体系时使用的数学转化方法；

投影参数(Projection Parameters)：对数据集进行投影操作时用于控制投影误差、变形实际分布的参考特征；

地图(Map)：空间现象的空间表征，通常以平面图形表示；

现象(Phenomenon)：事实、发生的事件、状态等；

分辨率(Resolution)：由涉及到或使用的测量工具或分析方法能区分开的两个独立测量或计算的值的最小差值；

质量(Quality)：数据符合一定使用要求的基本或独特的性质；

详述(Explicit)：由一对数或三个数分别直接描述水平位置和三维位置的方法；

介质(Media)：用于记录、存储或传递数据的物理设备。

4．1．3 空间数据元数据的标准

同物理、化学等学科使用的数据结构类型相比，空间数据是一种结构比较复杂的数据类型。它涉及到对于空间特征的描述，也涉及到对于属性特征及其它们之间关系的描述，所以空间数据元数据标准的建立是项复杂的工作；并且由于种种原因，某些数据组织或数据用户

开发出来的空间数据元数据标准很难为地学界所广泛接受。但空间数据元数据标准的建立是空间数据标准化的前提和保证，只有建立起规范的空间数据元数据才能有效利用空间数据。目前，针对空间数据元数据，已经形成了一些区域性的或部门性的标准*。

4．2 空间数据元数据的应用

4．2．1 帮助用户获取数据

通过元数据，用户可对空间数据库进行浏览、检索和研究等。一个完整的地学数据库除应提供空间数据和属性数据外，还应提供丰富的引导信息，以及由纯数据得到的分析、综述和索引等。通过这些信息用户可以明白一系列问题，如“这些数据是什么数据？”，“这个数据库是否有用？”等。

4．2．2 空间数据质量控制

无论是统计数据还是空间数据都存在数据精确问题，影响空间数据精度的原因主要有两个方面：一是源数据的精度；一是数据加工处理工程中精度质量的控制情况。空间数据质量控制内容包括：（1）有准确定义的数据字典，以说明数据的组成，各部分的名称，表征的内容等；（2）保证数据逻辑科学地集成，如植被数据库中不同亚类的区域组合成大类区，这要求数据按一定逻辑关系有效的组合；（3）有足够的说明数据来源、数据的加工处理工程、数据解译的信息。这些要求可通过元数据来实现，这类元数据的获取往往由地学和计算机领域的工作者来完成。数据逻辑关系在数据中的表达要由地学工作者来设计，空间数据库的编码要求一定的地学基础，数据质量的控制和提高要有数据输入、数据查错、数据处理专业背景知识的工作人员，而数据再生产要由计算机基础较好的人员来实现。所有这方面的元数据，按一定的组织结构集成到数据库中构成数据库的元数据信息系统来实现上述功能。

4．2．3 在数据集成中的应用

数据集层次的元数据记录了数据格式、空间坐标体系、数据的表达形式、数据类型等信息；系统层次和应用层次的元数据则记录了数据使用软硬件环境、数据使用规范、数据标准等信息。这些信息在数据集成的一系列处理中，如数据空间匹配、属性一致化处理、数据在各平台之间的转换使用等是必要的。这些信息能够使系统有效地控制系统中的数据流。

4．3 在地理信息系统中使用元数据的原因

在地理信息系统中使用元数据，有利于空间数据的管理共享，有利于实现一些特定功能，对于地理信息系统软件的开发，可以提高开发的效率和质量。

---

* 空间数据元数据标准的具体介绍见“地理信息系统标准”一章。

4．3．1 性能上的原因

1）完整性（Completeness）

面向对象的地理信息系统和空间数据库的目标之一，是把事物的有关数据都表示为类的形式，而这些类也包括类自身，即复杂的“类的类”结构。这就要求有支持类与类之间相互印证和操作的机制，而元数据可以帮助这个机制的实现。

2）可扩展性(Extensibility)

有意地延伸一种计算机语言或者数据库特征的语义是很有用的，如把跟踪或引擎信息的生成结果添加到操作请求中，通过动态改变元数据信息可以实现这种功能。

3）特殊性(Specialization)

继承机制是靠动态连接操作请求和操作体来实现的，语言及数据库以结构化和语义信息的相关上下文（Context）方式把操作请求传递给操作体，而这些信息可以通过元数据表达。

4）安全性(Safety)

分类完好的语言和数据库都支持动态类型检测，类的信息表示为元数据，这样在系统运行时，可以被类检测者访问。

4．3．2 功能上的原因

1）查错功能(Debugging)

在查错时使用元数据信息，有助于检测可运行应用系统的解释和修改状态。

2）浏览功能(Browsing)

为数据的控制类开发浏览器时，为显示数据，要求能解释数据的结构，而这些信息是以元数据来表达的。

3）程序生成(Program Generation)

如果允许访问元数据，则可以利用关于结构的信息自动生成程序，如数据库查询的优化处理和远程过程调用残体（或“桩”，stub）生成。

4．4 空间数据元数据的获取与管理

4．4．1 空间数据元数据的获取

空间数据元数据的获取是个较复杂的过程，相对于基础数据的形成时间，它的获取可分为三个阶段：数据收集前、数据收集中和数据收集后。对于模型元数据，这三个阶段分别是模型形成前、模型形成中和模型形成后。

第一阶段的元数据是根据要建设的数据库的内容而设计的元数据，内容包括：普通元数据、专指性元数据；第二阶段的元数据随数据的形成同步产生；第三阶段的元数据是在上述数据收集到以后，根据需要产生的，包括数据处理过程描述、数据利用情况、数据质量评估、浏览文件的形成、拓扑关系、影像数据的指标体及指标、数据集大小、数据存放路径等。

空间数据元数据的获取方法主要有五种：键盘输入、关联表、测量法、计算法和推理法。键盘输入一般工作量大且易出错；关联表方法是通过公共项（字段）从已存在的元数据或数据中获取有关的；测量法容易使用且出错较少，如用全球定位系统测量数据空间点的位置等；

计算方法指由其它元数据或数据计算得到的元数据，如水平位置可由仪器设置及时间计算得到；推理方法指根据数据的特征获取元数据。在元数据获取的不同阶段，使用的方法也有差异。在第一阶段主要是键入方法和关联表方法；第二阶段主要采样测量方法；第三阶段主要方法是计算和参考方法。

4．4．2 空间数据元数据的管理

空间数据元数据的理论和方法涉及到数据库和元数据两方面。由于元数据的内容、形式的差异，元数据的管理与数据涉及的领域有关，它是通过建立在不同数据领域基础上的元数据信息系统实现的。在元数据管理信息系统中，物理层存放数据与元数据，该层由一些软件通过一定的逻辑关系与逻辑层关联起来。在概念层中用描述语言及模型定义了许多概念，如实体名称、别名等。通过这些概念及其限制特征，经过与逻辑层关联可获取、更新物理层的元数据及数据。

4．5 元数据存储和功能实现

元数据系统用于数据库的管理，可以避免数据的重复存储，通过元数据建立的逻辑数据索引可以高效查询检索分布式数据库中任何物理存储的数据。减少数据用户查询数据库及获取数据的时间，从而减低数据库的费用。数据库的建设和管理费用是数据库整体性能的反映，通过元数据可以实现数据库的设计和系统资源的利用方面开支的合理分配，数据库许多功能（如数据库检索、数据转换、数据分析等）的实现是靠系统资源的开发来实现的，因而这类元数据的开发和利用将大大地增强数据库的功能并降低数据库的建设费用。

伴随着人类对数字地理信息重要性认识的加深，元数据标准化这一问题便逐渐成为共享地学信息的热点，而要研究元数据体系，则首先要对元数据的理论基础有一个正确的分析。事实上元数据标准依赖于信息共享标准的理论，它与自然科学中的许多学科都有交叉，几乎涉及数理化天地生中的所有方面，并依赖于现代科技的发展。计算机是它的基础平台，网络是它的通讯基础，没有数学模型和对各学科的综合认识，也就谈不上用遥感等技术研究地球机理。因此，从宏观角度来看，地理信息标准化涉及许多领域，似乎它的理论也枚不胜举；但从微观角度来考虑，数字地理信息所研究的共享体系理论则主要包括地理信息的模型建立表示理论、空间参照系理论、质量体系理论以及计算机通讯技术等方面的理论，它们是数据共享体系的基础。当然，其它能够促使地理信息共享的理论也将成为基于数字地球的元数据体系的有力支柱。

地理信息的质量体系

（1）质量评价过程

地理空间数据的质量对数据生产者和用户来说都是一个非常重要的考虑因子，它可以使数据生产者正确描述他们的数据集符合生产规范的程度，也是用户决定数据集是否符合他们应用目的的依据。因此，探索数据质量的理论问题便成为地理空间数据标准化的重要组成部分。在元数据标准中，质量信息主要在标识信息、数据质量信息及其数据继承关系等元数据部分中出现，其所涉及的主要元素有数据集的完备性、逻辑一致性、位置精度、时间精度、专题精度等，而每一元素又有各自的子元素。由于用户需要不同层次的数据质量，有些用户需要高精度的信息，而另一些用户则有较低层次的精度便可以满足他们的需求，这样对数据集的质量标准就有不同的评价依据。但作为质量中的几何精度评价，则通过一定的计算公式和相应的精度指标，可以获得数据集的几何精度。

（2）多尺度评价依据

在质量评定过程中，一般来说，数据的精度或准确度越高越好，但在实际应用中却不能不分对象一概而论。事实上有的数据在应用中的意义很大（如大地控制点等），其本身精度也可以达到很高，因此对这些数据的精度要求也就很高；而另一些数据本身的精度不可能很高，如不同土壤类型的面积，由于它们之间的界限是模糊的，所以面积也是相对的，因此精度要求不可能很高；有的数据的精度可以达到很高，但需要花费很多的人力、物力和时间，而生产上或应用上又不一定要求很高。因此，在实际应用中应根据具体需要来评定数据的质量。地球是一个复杂的系统，不少物体具有不确定性或模糊性特征。有些物体本身就没有明确的界限，它们是逐渐过渡的，在由量变到质变过程中，难以确定其边界线；有些虽然有明确的定义，但很难操作；有些数据是动态的，甚至是瞬间的。通过上面的分析，应该在地理信息描述中辨证地分析精度问题，既要追求很高的数据精度，又要避免“冗余”精度，以避免造成精度浪费。

（3）数据的实效性与唯一性

地球系统的数据，有些具有明显的时效（时间）特征，有的则对时间的反应比较迟缓，例如土地利用图的时效性来说，随地区的差异在时间上有明显的变化；而相对来说地质图、地形图则没有明显的时效性。

在图形的时效性上，一般来说具有动态特征的数据，它们的时间有效性较短，相反则长。但从研究历史变化或发展过程来说，任何时间的数据都是有用的。所以不同时效性的数据集将根据其作用的不同来确定其重要性，而这些因素在元数据体系中都应有所反应。

另外，地球系统的数据有可派生数据和不可派生数据之分，而在数据集描述中应避免派生数据。如在气象和水文数据中，每天的降雨量是基本数据，或不可派生数据，而月平均降雨等都是派生的。因此，在元数据描述中，应只限于基础数据，而不应包括可派生数据。这样，便需要有专门的元素来描述数据集的这些特征，并需要对不同的计算公式等加以阐述。

（4）数据精度的测试与报告

对于用户和数据生产者来说，所关心的数据质量是有一定联系的。数据集生产者必须使所生产的数据集满足制图规范，而用户则根据数据集的质量信息确定该数据集是否满足他们的应用需求。因此数据集生产者提供的数据集信息应是用户所关心的信息。因此，在数据集报告中应包括相应的精度测试方法及其测试结果等内容。

万事开头难。

中国俗语

第六章 空间数据获取

导读：空间数据获取是地理信息系统建设首先要进行的任务，它可以有多种实现方式包括数据转换、遥感数据处理以及数字测量等等，其中已有地图的数字化录入，是目前被广泛采用的手段，也是最耗费人力资源的工作。在 GIS 中，录入的内容包括空间信息和非空间信息，前者是录入的主体。目前，空间信息的录入主要有两种方式，即手扶跟踪数字化和扫描矢量化，本章具体介绍了两种方式，以及相关的算法，如曲线近似拟合，栅格图形细化跟踪等。

在图形数据录入完毕后，需要进行各种处理，包括坐标变换、拼接等等，其中最重要的是建立拓扑关系。在拓扑建立过程中，需要先对各种错误修改，本章描述了各种具体的错误情形，最后则介绍了多边形自动拓扑生成算法。

1．地图数字化

1．1 概述

无论古代、现代还是将来，地图总是重要的信息形式。在计算机图形图像技术应用于地图制作之前，各部门一般都使用纸质地图或工程图纸，纸图在查阅、计算距离和标注地名符号等方面都是人工操作；另外不能对一幅纸地图进行修改、缩小比例和分层读图，至于地图局部放大只能借助于放大镜，被放大的区域很小且操作不方便。

随着技术的发展，人们对地图的要求进一步提高。由于传统纸地图效率、速度和精度很低，因此难以适应现代和未来科技发展。而通过 GIS 工具，可以把纸地图经过一系列处理而转换成可以在屏幕上显示的电子化地图，可以满足人们使用地图的新的要求。为了讨论问题方便，将矢量电子地图定义如下：

当纸地图经过计算机图形图像系统光——电转换量化为点阵数字图像，经图像处理和曲线矢量化，或者直接进行手扶跟踪数字化后，生成可以为地理信息系统显示、修改、标注、漫游、计算、管理和打印的矢量地图数据文件，这种与纸地图相对应的计算机数据文件称为矢量化电子地图。这种地图工作时需要有应用软件和硬件系统的支撑。对矢量化地图的操作是以人机交互方式，通过 GIS 应用软件对硬件设备的控制来实现的。

在硬件系统及相应的地理信息系统软件支持下，矢量电子地图与纸地图相比有如下优点：

．计算距离和标注地名符号快速准确；

．可对地图局部放大、全图缩小和移动显示、漫游功能很强；

．分层显示地图（当对地图上各种信息分不同层归类存放后，则可以显示某些层，关闭不显示的层）；

．可以以图元为单位进行信息编辑修改，人机交互画线标注符号文字，删除地图上多余的信息；

．可以通过计算机网络进行电子地图传递，提供信息共享，传递的速度快，保密性强；

．如果能有效解决地图符号自动分割和识别问题，则能实现地图的智能矢量化。这里智能化是指自动矢量化和自动标注符号，最佳路径优化选择和自动跟踪目标等。

矢量电子地图与点阵地图图像相比有如下优点：

．相同信息量下前者的文件相对要小得多，图越复杂表现越明显；

．前者可以以图元为单位进行信息编缉修改删除，人机交互画线标注符号文字；后者只能以像素为基本单位（如矩形图像块）进行拷贝，移动和删除，即它的编辑功能很差；

．前者可对所有图元分层显示，后者只能做到对整图某区域（矩形区）的开窗显示控制。

1．2 地图数据类型

为地图自动制图目的所搜集的上述资料，要适应计算机处理的需要，必须同时考虑通过这些资料，最后能获取各种不同性质的地图数据。

为此，可以定义地图数据是载荷地理信息的数字集合。它们应指被输入到某种计算机系统中进行处理，必要时可重新生成人的视觉可以感受的地理信息。在这一层意义上，地图数据可以分为：

1）空间数据（或图形数据）

空间数据是构成地图内容要素的几何图形，例如地图上的井、山峰、灯塔、河流、道路、等值线、湖泊、森林界线和土壤类型界线等。为表示这些要素在二维平面上空间图形的定位特征，常用一对平面直角坐标（*X, Y*）来表示，这种地图数据称为矢量数据；或用其通过栅格单元的左下角坐标（行和列）来表示，称此为栅格数据。

地图要素图形大致可以分为点、线、面三种基本类型。

面——可由环绕它们的线表示；也可以由其区域内的点表示；

线——可离散化成为点的集合；

而点则能用一对平面坐标系中的坐标来确定。

在实际应用中，仅有坐标数据是不够的，必须依照不同地图要素的意义，通过特别的编码加以区别。

2）语义数据(属性数据)

语义数据又称为非几何数据，包括定性数据和定量数据。定性数据用来描述要素的分类或对要素进行标名。定量数据是说明要素的性质、特征或强度的，例如距离、面积、人口、产量、收人、流速，以及温度和高程等.

对于语义数据，基本上是一个地理编码问题。以科学的分类分级系统为基础，对地理环境中各基本实体及其联系进行编码，以便唯一地对某一系统中所有地图要素进行认别和处理，这种功能不仅在用户环境中而且在计算机系统内也是十分重要的。

当前，要为满足各方面全部需求而建立一种通用的编码是不可能的。因此，最现实的是研制多样性编码满足各部门的不同需要，同时兼顾符合建立数据库的规定。一般来说编码的一些基本要求包括：

（2．1）要素类别——如地名，实体类型及等级等；

（2．2）要素特征（属性）——每一要素可具有与它有联系的大量的属性值，它承载关于要素特征的信息；

（2．3）作用范围的描述——例如，一个区域的土壤类型；

（2．4）地理定义——在某些情况下编码可以是一个实体集合的间接参考，而集合中的每一个实体也会被它自身的地理代码所说明，如水系、地质层的编码。

在 GIS 数据录入过程中，空间数据的录入更为重要，因为处理空间数据和空间信息正是 GIS 的核心功能。与属性数据录入相比，空间数据录入的工作量更大，并且需要 GIS 软件工具的支持（利用通用的数据库软件即可以进行属性数据录入）。

1．3 数字化仪数字化

1．3．1 手扶跟踪数字化

尽管手扶跟踪数字化（Manual Digitising）工作量非常繁重，但是它仍然是目前最为广泛采用的将已有地图数字化的手段。

利用手扶跟踪数字化仪可以输入点地物、线地物以及多边形边界的坐标。其具体的输入方式与地理信息系统软件的实现有关，另外一些 GIS 系统也支持用数字化仪输入非空间信息，如等高线的高度，地物的编码数值等等。

手扶跟踪数字化仪的通讯和参数：

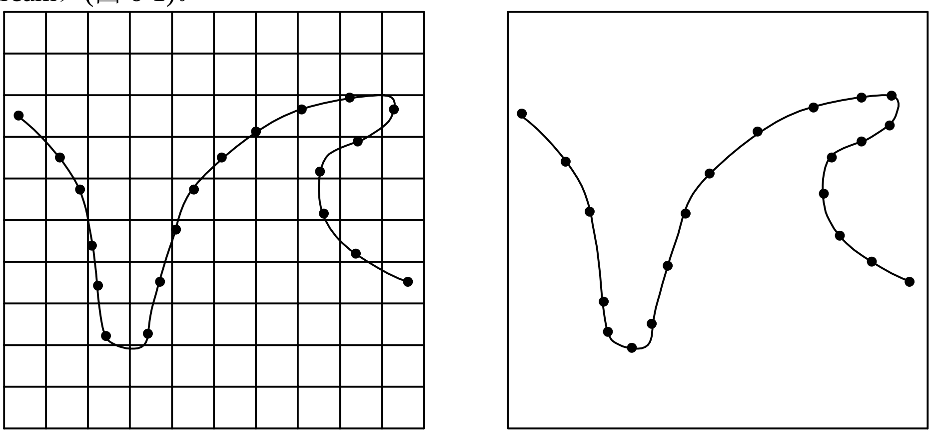
手扶跟踪数字化仪是通过 RS-232（串口）接口与计算机进行连接的，为了能够进行正确的数据发送和接收，需要进行通讯参数的设置，包括波特率、数据位、校验位、停止位等等。此外，数字化仪还包括坐标原点、分辨率、采点方式、数据格式等参数。数字化仪的参数通常可以利用数字化板上的开关和菜单确定。为了保证数据录入的正确，必须设置数字化软件的参数与数字化仪的一致。

在进行数字化之前，首先要确定需要数字化哪些信息，在目前由于大多数 GIS 软件对空间数据采用分层管理，所以要确定输入哪些图层，以及每个图层包含的具体内容。另外，由于数字化过程不可能一次完成，在两次输入之间地图的位置可能相对于数字化板发生错动，这样前后两次录入的坐标就会偏移或旋转。解决该问题的办法就是，在每次录入之前，先输入至少三个定位点（Tick Marks），或称为注册点（Register Points），这些点相对于地图的位置是固定的，这样两次输入的内容就可以根据定位点坐标之间的关系进行匹配。

通常，数字化仪采用两种数字化方式，即点方式（Point Mode）和流方式（Stream Mode），点方式是当录入人员按下游标(Puck)的按键时，向计算机发送一个点的坐标。输入点状地物要素时必须使用点输入方式；而线和多边形地物的录入可以使用点方式，在输入时，输入者可以有选择地输入曲线上的采样点，而采样点必须能够反映曲线的特征。

流方式录入能够加快线或多边形地物的录入速度，在录入过程中，当录入人员沿着曲线移动游标时，能够自动记录经过点的坐标。采用流方式录入曲线时，往往采集点的数目要多于点方式，造成数据量过大，一个解决的方案是对记录的点进行实时采样，即尽管系统接收到了点的坐标，但是可以根据采样原则确定是否记录该点*。

目前大多数系统采取两种采样原则，即距离流方式（Distance Stream）和时间流方式（Time Stream）(图 6-1)。



(a) (b)

图 6-1：距离流方式和时间流方式

(a)距离流方式是当前接收的点与上一点距离超过一定阈值，才记录该点；

* 也可以在整个曲线录入完毕后，使用下面描述曲线离散化算法去掉一些非特征点。

(b)采用时间流方式时，按照一定时间间隔对接收的点进行采样。

采用时间流方式录入时，一个优点是当录入曲线比较平滑时，录入人员往往移动游标比较快，这样记录点的数目少；而曲线比较弯曲时，游标移动较慢，记录点的数目就多。而采用距离流方式时，容易遗漏曲线拐点，从而使曲线形状失真。所以在保证曲线的形状方面，时间流方式要优于距离流方式。

在实际的录入过程中，可以根据不同的录入对象选择不同的录入方式。例如，当录入地块图时，由于其边界多为直线，并且点的数据较少，可以采用点方式录入；录入交通线时，因为要保证某些特征点位置的准确性，也可以使用点方式；而等高线的录入由于数据量大，使用流方式可以加快录入速度。

其它的矢量数据录入方式：

尽管手扶跟踪数字化是目前最主要的矢量格式地理数据的录入手段，但是在某些场合下，也可以采用其它的数据输入方式：

1）其它数据转换：

其它格式数据的转换包括三种情形：

（1．1）其它矢量格式数据（往往是由其它 GIS 软件制订）的转换；

（2．2）坐标数据，往往表现为关系数据库表的形式（表 6-1）；

（2．3）位置描述信息，以关系数据表形式存取，同样可以转换为不太精确的坐标数据（表 6-2）。

表 6-1：测站信息表（部分）

测站编码	经度	纬度
68013344	107.2	29.8
68026785	115.5	30.2

表 6-2：企业员工信息表（部分）

姓名	住址
张三	北京市海淀区
李四	河北省石家庄市

2）键盘录入：

对于数据量较小、并且已知地物精确坐标的情况下，可以采用键盘录入。此外键盘录入也是录入属性数据的主要手段。

3）鼠标录入：

如果不愿意手扶跟踪数字化，而扫描矢量化又难以识别地物时，可以使用鼠标录入，通常是将地图扫描后，作为底图显示在屏幕上，用鼠标参照底图进行采点。由于鼠标定位不如数字化仪精确，所以一般用于输入一些示意图。

4）其它定点测量设备：

目前 GPS 已经成为流行的定位导航设备，它同样可以为 GIS 提供矢量格式的坐标数据。其它的定点测量设备包括平板测图仪等，使用这些设备，最重要的是要考虑精度问题。

1．3．2 曲线离散化算法

在数字化过程中，需要对曲线进行采样简化，即在曲线上取有限个点，将其变为折线，并且能够在一定程度上保持原有的形状。下面介绍 Douglas-Peucker 算法（图 6-2）。

1）在曲线首尾两点 *A*、*B* 之间连接一条直线段 *AB*，该直线称为曲线的弦；

2）得到曲线上离该直线段距离最大的点 *C*，并计算其与 *AB* 的距离 d；

3）比较该距离与预先给定阈值 ε 的大小，如果小于 ε，则将该直线段作为曲线的近似，该段曲线处理完毕；

4）如果距离大于阈值，则用 *C* 将曲线分为两段 *AC* 和 *BC*，并分别对两段曲线进行 1-3 步的处理。

5）当所有曲线都处理完毕后，依次连接各个分割点形成的折线，即可以作为曲线的近似。

很明显，该算法是一个递归算法。

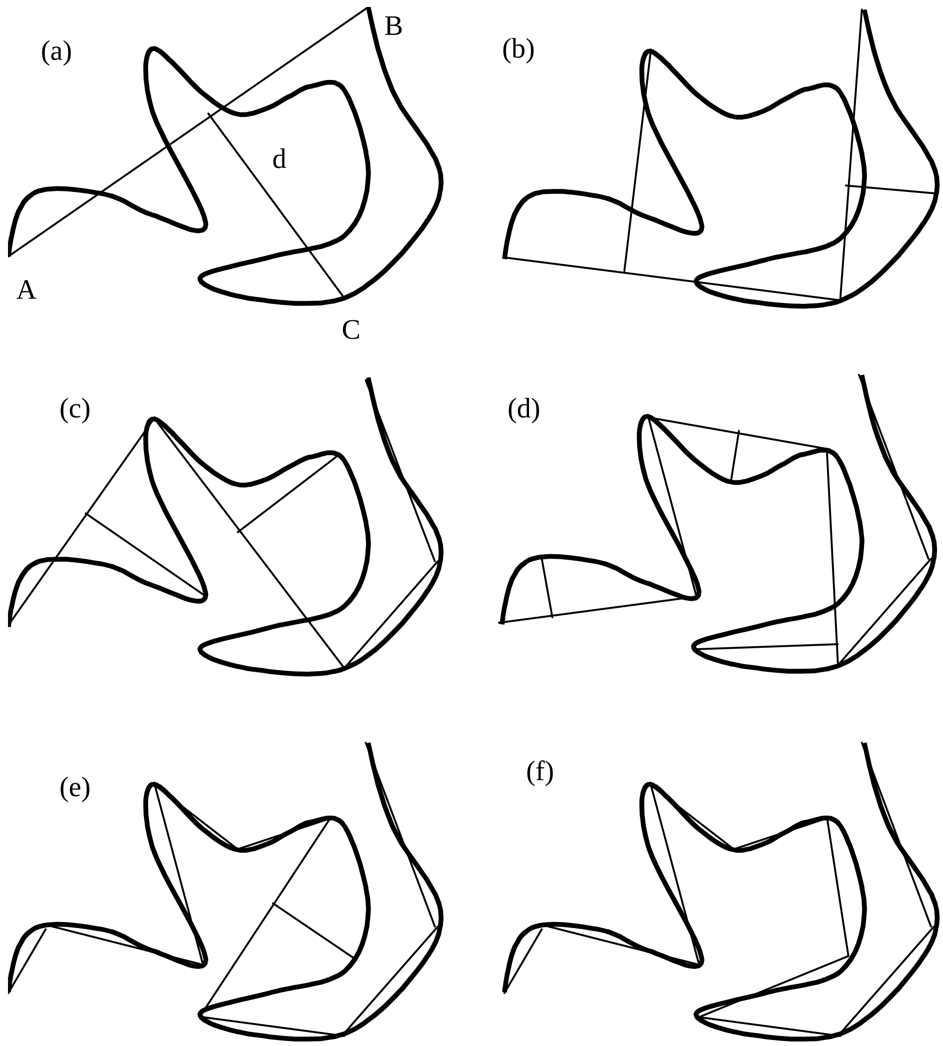


图 6-2：曲线的离散算法

1．4 扫描矢量化及常用算法

1．4．1 扫描矢量化以及处理流程

随着计算机软件和硬件更加便宜，并且提供了更多的功能，空间数据获取成本成为 GIS 项目中最主要的成分。由于手扶跟踪数字化需要大量的人工操作，使得它成为以数字为主体的应用项目瓶颈。扫描技术的出现无疑为空间数据录入提供了有力的工具。

常见的地图扫描处理的过程如图 6-3 所示。由于扫描仪扫描幅面一般小于地图幅面，因此大的纸地图需先分块扫描，然后进行相邻图对接；当显示终端分辨率及内存有限时，拼接后的数字地图还要裁剪成若干个归一化矩形块，对每个矩形块进行矢量化（Vectorization）处理后生成便于编辑处理的矢量地图，最后把这些矢量化的矩形图块合成为一个完整的矢量电子地图，并进行修改、标注、计算和漫游等编辑处理。

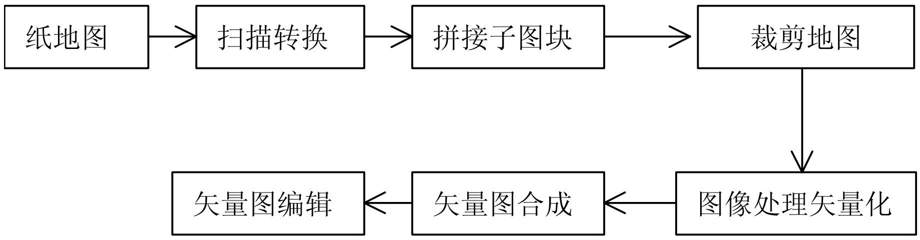


图 6-3：地图信息处理流程图

在扫描后处理中，需要进行栅格转矢量的运算，一般称为扫描矢量化过程。扫描矢量化可以自动进行，但是扫描地图中包含多种信息，系统难以自动识别分辨（例如，在一幅地形图中，有等高线、道路、河流等多种线地物，尽管不同地物有不同的线型、颜色，但是对于计算机系统而言，仍然难以对它们进行自动区分），这使得完全自动矢量化的结果不那么“可靠”，所以在实际应用中，常常采用交互跟踪矢量化，或者称为半自动矢量化。

将栅格图像转换为矢量地图一般需要以下一系列步骤[Musavi 1988]：

1）图像二值化（Threshold）

图像二值化用于从原始扫描图像计算得到黑白二值图像（Binary Image），通常将图像上的白色区域的栅格点赋值为 0；而黑色区域为 1，黑色区域对应了要矢量化提取的地物，又称为前景。

2）平滑（Smooth）

图像平滑用于去除图像中的随机噪声，通常表现为斑点。

3）细化

细化将一条线细化为只有一个像素宽，细化是矢量化过程中的重要步骤，也是矢量化的基础。

4）链式编码

链式编码将细化后的图像转换成为点链的集合，其中每个点链对应于一条弧段。

5）矢量线提取

将每个点链转化成为一条矢量线。每条线由一系列点组成，点的数目取决于线的弯曲程度和要求的精度。

除了上述五个步骤以外，还需要一些处理以方便图像矢量化过程，如图像拼接和剪裁等等，下面对这些操作以及相关算法进行描述。

1．4．2 图像拼接/裁剪

1）图像拼接

以两相邻地图图像的部分重叠区为基础，把它们合成为一幅整图的过程叫做图像拼接，分上下拼接和左右拼接。以左右拼接为例，取左图右边缘一个矩形区域 A，取右图左边缘一个矩形区域 B，如果 A 和 B 有一定的重叠区，可以利用计算机实现自动的匹配，其拼接算法如下：

（1．1）由 A 中右侧边缘从右至左依次取若干个列 L1（0），L1（1），…，L1（n），以各列内像素灰度的长度序列为特征向量，分别求出以上各列的特征向量 *V1（0），V1（1），…，V1（n）*。

（2．2）自动拼接 即由 B 中左侧从左至右依次取若干列 L2（n），L2（n-1），…，L2（0），以各列内像素灰度的长度序列为特征向量，分别求出以上各列的特征向量 *V2（n）：V2（n-1），…，V2（0）*。若向量序列 *[V1（0）；V1（1），…，V1（n）]* 和 *[V2（0），V2（1），…，V2（n）]* 匹配，则转向步骤 4。

（2．3）人工拼接 即固定 A，通过人机交互控制 B 以一定步长上下左右移动，直到 A 和 B 重叠区对齐为止。

（2．4）根据步骤 2 匹配情况或步骤 3 的偏移情况对两相邻地图图像进行修正和合成。

2）图像裁剪

把一幅图像裁成两两相邻的规则图块的过程称为地图裁剪。图像裁剪非常简单，实际应用中，可以根据不同的硬件配置确定采用和不采用图像裁剪技术。

1．4．3 图像细化预处理二值图像平滑

在将地图扫描或摄像输入时，由于线不光滑以及扫描、摄像系统分辨率的限制，使得一些曲线目标带来多余的小分支（即毛刺噪声）；此外，还有孔洞和凹陷噪声，如图 6-4 所示。如果不在细化前去除这几种噪声，就会造成细化误差和失真，这样会最终影响地图跟踪和矢量化。曲线目标越宽，提取骨架和去除轮廓所需的次数也越多，因此噪声影响也越大。

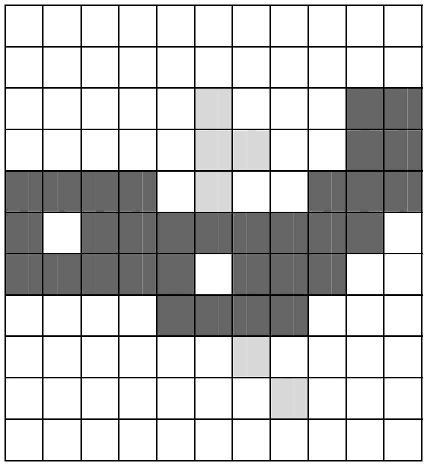


图 6-4：扫描图像的“毛刺”和“凹陷孔洞”

为了去除毛刺噪声的影响，可以采用如图 5 所示的 3×3 模板进行处理。处理的过程是：按点阵格式扫描图像上每一像素，只要图像相应区域与图 6-5 中的模板（包括其三次 90^0旋转所形成的模板）匹配，则判定为毛刺，对应于模板中心的像素数值变为0。根据需要可进行多次这种匹配运算。

0	0	0
0	1	0
X	X	X

图 6-5：去毛刺模板，X 为任意数值

为了去除孔洞及凹陷噪声，我们采用如图 6-6 所示的模板进行处理，只要图像对应区域与该模板（包括其三次 90^0旋转）匹配，则区域中心点数值变为 1。

X	1	X
1	0	1
X	X	X

图 6-6：去孔洞凹陷模板

总之，通过以上两种平滑处理，基本上消除了毛刺和孔洞凹陷噪声的影响，为进一步进行细化处理打下了基础。

1．4．4 图像细化（Thinning）

细化算法对二值图像进行处理，得到细化后的图像。线细化是处理包含线状地物二值图

像的一种重要技术，在地图扫描处理中，由于地图上主要信息是不同粗细和不同形状的线，必须首先进行线细化，以准确、有效地提取这些线信息，并进一步完成跟踪矢量化。

线细化，就是不断去除曲线上不影响连通性的轮廓像素的过程，对细化的一般要求是：

- 保证细化后曲线的连通性
- 细化结果是原曲线的中心线
- 保留细线端点

根据各种不同的应用，目前已经提出了许多线细化算法，如内接圆法、经典算法、异步算法、快速并行算法及并行八边算法等，不同的算法在处理速度和效果上各有其特点。

下面介绍一个常用的细化算法，其它算法基本是此算法的改进。

首先介绍几个相关的概念和符号。对于二值栅格图像中每个像素点 p，以及该像素直接相邻的 8 个像素点（图 6-7），令：

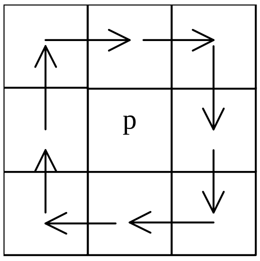


图 6-7：像素周围的 8 个直接相邻像素

1）$N(p)$为 p 的邻点的数值的和；

2）图像像素联接数 $T(p)$，如果旋转着看像素周围的点，$T(p)$就是 p 周围 8 个点从 0 变成 1 的次数，它反映了像素邻点的联接的块数（图 6-8）。

3）p_W，p_E，p_S，p_N分别指像素左侧、右侧、下边、上边邻点的数值。

	1	

(a)T=0

1	1	1
1	1	1
1	1	1

(b)T=0

		1
	1	

(c)T=1

1	1	
1	1	
1		

(d)T=1

1	1	
1	1	
	1	

(e)T=2

1		1
	1	

(f)T=2

1	1	1
	1	
1	1	1

(g)T=2

1		1
	1	
	1	

(h)T=3

1		1
	1	
1		1

(i)T=4

图 6-8：像素联结数

算法步骤如下[Zhang-Suen，1984]：

1：对于栅格图像中的每个点 p，进行如下操作：

如果 $2\le N(p)\le 6$ 并且 $T(p)=1$ 并且 $p_N p_S p_E=0$ 并且 $p_W p_E p_S=0$

则标志 p 点；

2：将所有被标志的栅格点赋值为 0，如果没有被标志的点，则算法结束；
3：对于栅格图像中的每个点 p，进行如下操作：
　如果 $2\le N(p)\le 6$ 并且 $T(p)=1$ 并且 $p_N p_S p_W=0$ 并且 $p_W p_E p_N=0$
　则标志 p 点；
4：将所有被标志的栅格点赋值为 0，如果没有被标志的点，则算法结束；
5：转到第一步。
图 6-9 显示了采用该算法细化的过程和结果。

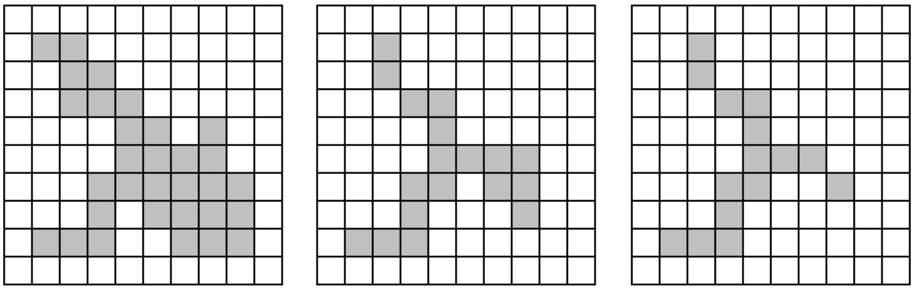


图 6-9：线状地物的细化

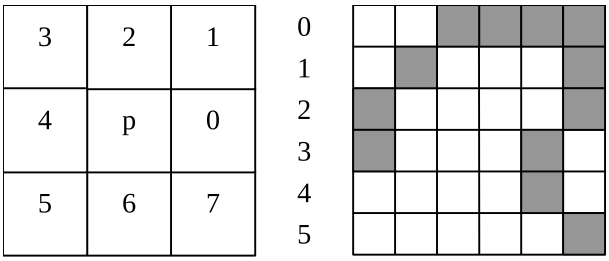
1．4．5 链码（弗里曼码）

链码是由弗里曼（Freeman）提出的用曲线出发点坐标和线的斜率来描述二值线图形的一种方法。图 6-10（a）所示是链码的八个方向及它们的序号。图 6-10（b）的细线的链码为(3,0)21100066567，其中（3,0）为起始点坐标，之后的数值序列描述了方向。

任意一条细线都可用链码序列表示为下式：

$C=a_1a_2...a_n，0\leqslant a_i\leqslant 7$

如果始点 a1 和终点 an 重合，则说明曲线是闭合的。



(a)　(b)

图 6-10：链码及其对细线的表示
(a)：链码的 8 个方向；(b)细线

1．4．5 矢量线生成

扫描矢量化的最后一步是生成矢量线，可以很方便地将链式编码的每一条链转换成为一条矢量线。自然地，弯曲的矢量线比直线需要更多的点，这还取决于要求的精度。在矢量线生成过程中，可以使用 Douglas-Peucher 算法。

目前，自动扫描矢量化还不是完全可靠的，为了提高其可靠性，需要在模式识别方面做出更多的研究。

2. 空间数据录入后的处理

2．1 图形坐标变换

在地图录入完毕后，经常需要进行投影变换，得到经纬度参照系下的地图。对各种投影进行坐标变换的原因主要是输入时地图是一种投影，而输出的地图产物是另外一种投影。进行投影变换有两种方式，一种是利用多项式拟合，类似于图像几何纠正；另一种是直接应用投影变换公式进行变换。

2．1．2 基本坐标变换

在投影变换过程中，有以下三种基本的操作：平移、旋转和缩放。

1）平移

平移是将图形的一部分或者整体移动到笛卡尔坐标系中另外的位置(图 6-11-a)，其变换公式如下：

$X' = X + T_x$

$Y' = Y + T_y$

2）缩放

缩放操作可以用于输出大小不同的图形（图 6-11-b），其公式为：

$X' = XS_x$

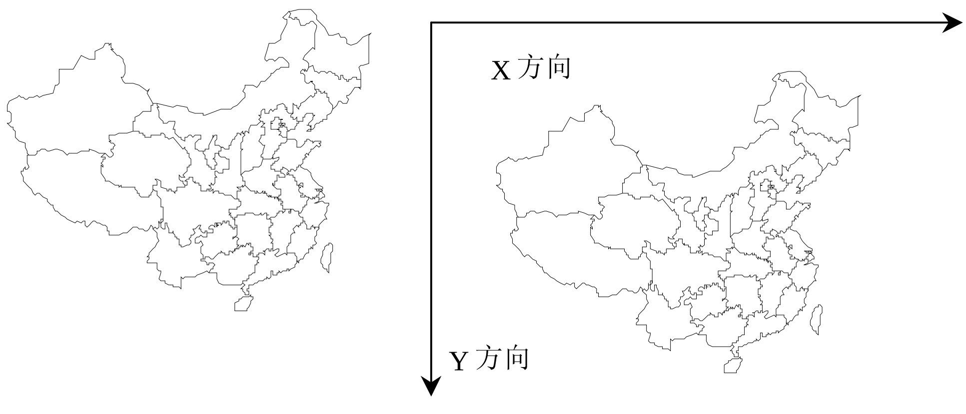
$Y' = YS_y$

3）旋转

在地图投影变换中，经常要应用旋转操作（图 6-11-c），实现旋转操作要用到三角函数，假定顺时针旋转角度为θ，其公式为*：

$X' = X\cos\theta + Y\sin\theta$

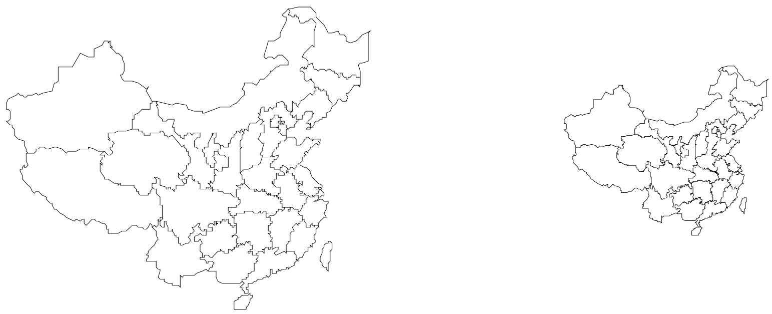
$Y' = -X\sin\theta + Y\cos\theta$



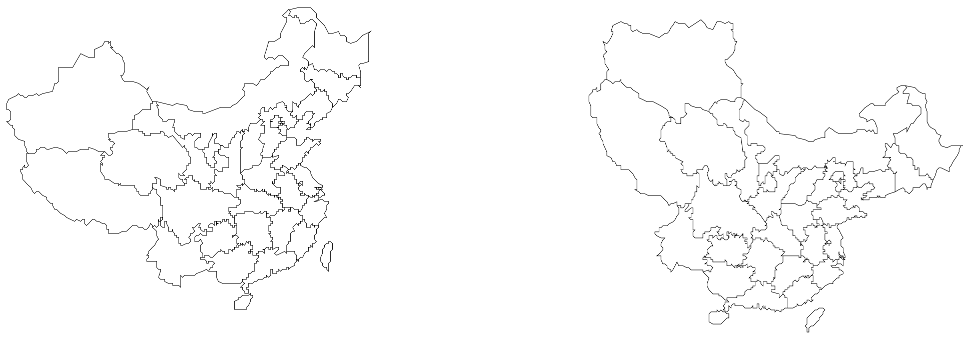
(a)平移

---

* 公式同“空间参照系统和地图投影”一章的坐标旋转公式，顺时针旋转图形相当于逆时针旋转坐标系。



(b)缩放



(c)图形旋转

图 6-11：图形坐标变换

2．1．2 仿射变换（Affine Tranformation）

如果综合考虑图形的平移、旋转和缩放，则其坐标变换式如下：

$$(X',Y') = \lambda \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \end{bmatrix}$$

上式是一个正交变换，其更为一般的形式是：

$$(X',Y') = \lambda \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \end{bmatrix}$$

后者被称为二维的仿射变换（Affine Transformation），仿射变换在不同的方向可以有不同的压缩和扩张，可以将球变为椭球，将正方形变为平行四边形（图 6-12）。

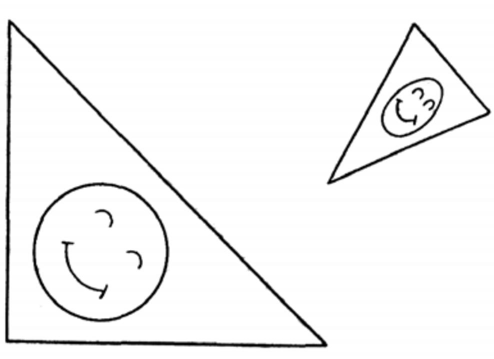


图 6-12：仿射变换

2．2 图形拼接

在对底图进行数字化以后，由于图幅比较大或者使用小型数字化仪时，难以将研究区域的底图以整幅的形式来完成，这是需要将整个图幅划分成几部分分别输入。在所有部分都输入完毕并进行拼接时，常常会有边界不一致的情况，需要进行边缘匹配处理（图 6-13）。边缘匹配处理，类似于下面提及的悬挂节点处理，可以由计算机自动完成，或者辅助以手工半自动完成。

除了图幅尺寸的原因，在 GIS 实际应用中，由于经常要输入标准分幅的地形图，也需要在输入后进行拼接处理，这时，一般需要先进行投影变换，通常的做法是从地形图使用的高斯——克吕格投影转换到经纬度坐标系中，然后再进行拼接。

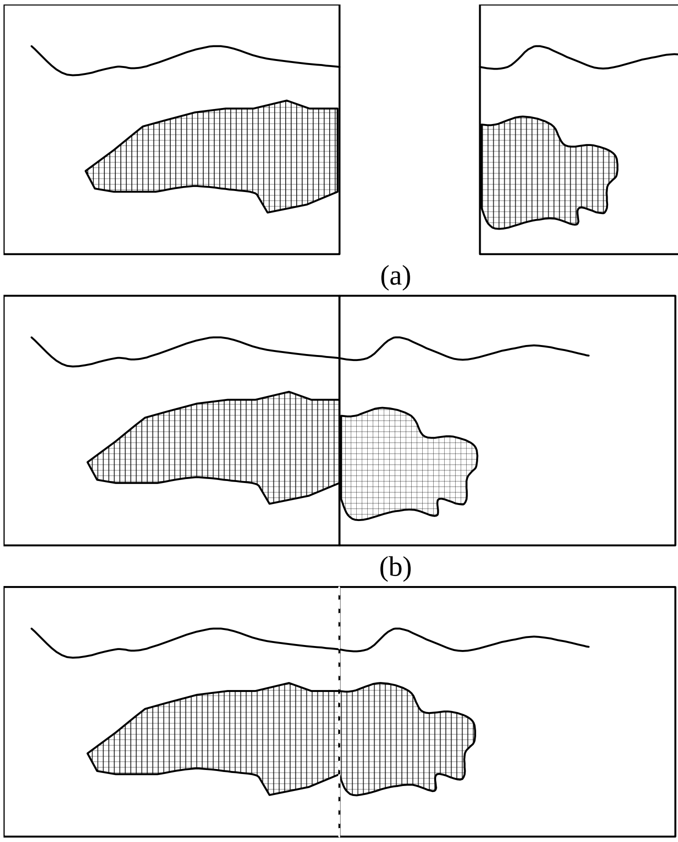


图 6-13：图幅拼接

(a)拼接前；（b）拼接中的边缘不匹配；(c)调整后的拼接结果

2．3 拓扑生成

在图形数字化——无论是手扶跟踪数字化还是扫描矢量化——完成后，对于大多数地图需要建立拓扑，以正确判别地物之间的拓扑关系。在 GIS 数据管理中，拓扑关系可以定义以下内容：

1）**区域**，如果多边形数据DIME数据模型，每个多边形可以用一组封闭的线*来表示，而不需要记录封闭线上的所有点，避免两次记录相邻多边形的公共边界，这样减少了数据冗余*。

---

* 通常称为弧段

* 与DIME数据模型相对应的是Spaghetti多边形编码方式，该编码方式对每个多边形都依次记录其边界上的

2）邻接性，另一种可以用拓扑描述的属性是多边形之间的相互邻接性。

3）连通性，连通性是指对弧段连接的判别，连通性的建立和表现是网络分析的基础。

2．3．1 图形修改

在建立拓扑关系的过程中，一些在数字化输入过程中的错误需要被改正，否则，建立的拓扑关系将不能正确地反映地物之间的关系。

ESRI 定义了以下判断录入图形是否正确的六个准则，可以帮助发现拓扑错误。

1）所有录入的实体都能够表现出来；

2）没有输入额外的实体；

3）所有的实体都在正确的位置上，并且其形状和大小正确；

4）所有具有连接关系的实体都已经连上；

5）所有的多边形都有且只有一个标志点以识别它们；

6）所有的实体都在边界之内

上述的准则，特别是第五和第六条，只是针对 ESRI 的 ARC/INFO 软件而言，其它的 GIS 软件由于具体实现的不同，可能会有差异。

由于地图数字化，特别是手扶跟踪数字化，是一件耗时、烦杂的人力劳动，在数字化过程中的错误几乎是不可避免的，造成数字化错误的具体原因包括：

1）遗漏某些实体；

2）某些实体重复录入，由于地图信息是二维分布的，并且信息量一般很大，所以要准确记录哪些实体已经录入，哪些实体尚未录入是困难的，这就容易造成重复录入和遗漏；

3）定位的不准确，数字化仪分辨率可以造成定位误差，但是人的因素是位置不准确的主要原因，如手扶跟踪数字化过程中手的抖动，两次录入之间图纸的移动都可以使位置不准确；更重要的，在手扶跟踪数字化过程中，难以实现完全精确的定位，例如在水系的录入中（图 6-14），将支流的终点恰好录入在干流上基本上是不可能的（图 6-14-a），更常见的是图（b）和（c）所示的两种情况。

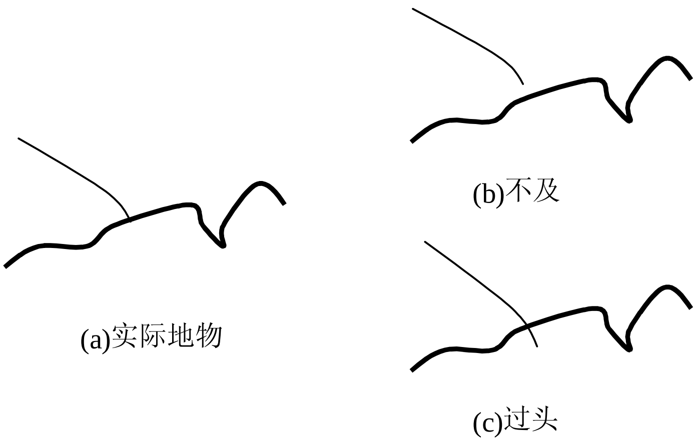


图 6-14：数字化错误——不及和过头

在数字化后的地图上，错误的具体表现形式有：

1）伪节点（Pseudo Node），伪节点使一条完整的线变成两段（图 15），造成伪节点的原因常常是没有一次录入完毕一条线。

顶点，显然公共边界被记录了两次。Spaghetti的意思是意大利细面条，用来形象的比喻这种数据结构。

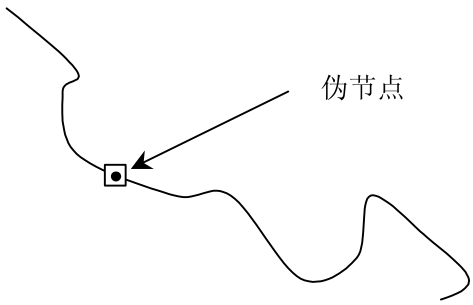
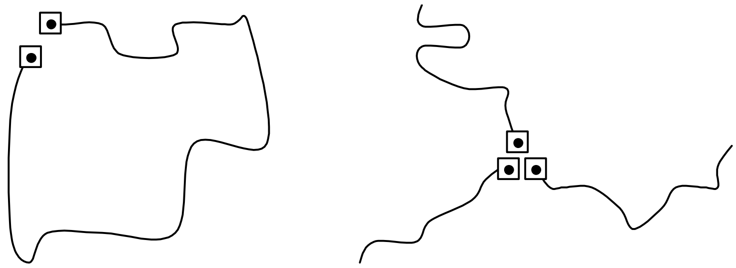


图 6-15：伪节点

2）悬挂节点（Dangling Node），如果一个节点只与一条线相连接，那么该节点称为悬挂节点，悬挂节点有多边形不封闭（图 6-16-a）、不及和过头（图 6-14-b，图 6-14-c），节点不重合（图 6-16-b）等几种情形。



(a)多边形不封闭　　　　　　　(b)节点不重合

图 6-16：悬挂节点的两种情形

3）“碎屑”多边形或“条带”多边形（Sliver Polygon）

条带多边形（图 6-17）一般由于重复录入引起，由于前后两次录入同一条线的位置不可能完全一致，造成了“碎屑”多边形。另外，由于用不同比例尺的地图进行数据更新，也可能产生“碎屑”多边形。

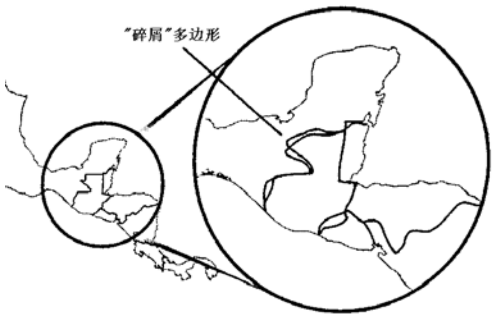
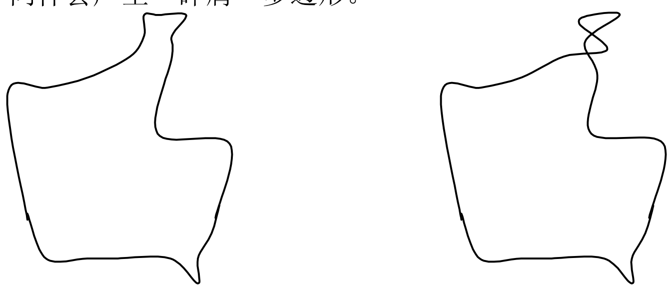


图 6-17：碎屑多边形

4）不正规的多边形（Weird Polygon）

不正规的多边形（图 6-18）是由于输入线时，点的次序倒置或者位置不准确引起的。在进行拓扑生成时，同样会产生“碎屑”多边形。



(a)正常多边形　　　　　　　(b)不正规多边形

图 6-18：不正规的多边形

上述的错误，一般会在建立拓扑的过程中发现，需要进行编辑修改。一些错误，如悬挂节点，可以在编辑的同时，由软件自动修改，通常的实现办法是设置一个“捕获距离”，当节点之间、或者节点与线之间的距离小于此数值后，即自动连接；而另外的错误需要进行手工编辑修改。

2．3．2 建立拓扑关系

在图形修改完毕之后，就意味着可以建立正确的拓扑关系，拓扑关系可以由计算机自动生成，目前大多数 GIS 软件也都提供了完善的拓扑功能；但是在某些情况下，需要对计算机创建的拓扑关系进行手工修改，典型的例子是网络连通性。

正如拓扑的定义所描述的，建立拓扑关系时只需要关注实体之间的连接、相邻关系，而节点的位置、弧段的具体形状等非拓扑属性则不影响拓扑的建立过程。

1）多边形拓扑关系的建立

如果使用 DIME 或者类似的编码模型，多边形拓扑关系的表达需要描述以下实体之间的关系：

- 多边形的组成弧段；
- 弧段左右两侧的多边形，弧段两端的节点；
- 节点相连的弧段。

多边形拓扑的建立过程实际上就是确定上述的关系。具体的拓扑建立过程与数据结构有关，但是其基本原理是一致的，下面简述多边形拓扑建立过程（图 6-19）。

图 6-19 中共有 4 个节点，以 A、B、C、D 表示；6 条弧段，用数字表示；以及 I、II、III 三个多边形（图 6-19-a）。首先定义以下概念：

- 由于弧段是有方向的，算法中将弧段A的起始节点称为首节点$N_s(A)$，而终止节点为尾节点$N_E(A)$；
- 考虑到弧段的方向性，沿弧段前进方向，将其相邻的多边形分别定义为左多边形和右多边形$P_L(A)$和$P_R(A)$。

在建立拓扑之前，首先将所有弧段的左右多边形（在实现中，可以用多边形的编码表示）都设置为空；然后对每个节点计算与其相连弧段的在连接处的角度，并进行排序（图 6-19-b）（注意，这个排序是循环的）。建立拓扑的算法如下：

（1）得到第一条弧段 A，并设置为当前弧段；

（2）判断$P_L(A)$和$P_R(A)$是否为空。如果都非空，转到第一步，当所有弧段处理完毕后，算法结束；

（3）如果左多边形为空，则创建一个新的多边形P，多边形的第一条弧段为当前弧段，并设置$P_L(A)=P$，设置搜寻起始节点为$N_s(A)$，搜寻当前节点为$N_E(A)$。如果右多边形为空，则创建一个新的多边形P，多边形的第一条弧段为当前弧段，并设置$P_R(A)=P$，设置搜寻起始节点$N_0=N_E(A)$，搜寻当前节点$N_C=N_S(A)$。

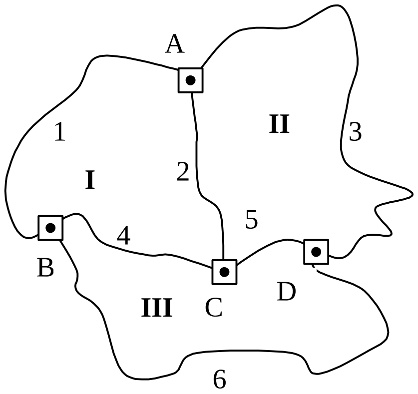
（4）判断N_0和N_C是否相等，如果是，则多边形所有弧段都已经找到，转到第一步。

（5）检查与当前节点相连接的、已经排列好的弧段序列，将当前弧段的下一条弧段 A' 作为多边形的第二条弧段。

（6）如果$N_C=N_S(A')$，设置$P_L(A')=P$，$N_C=N_E(A)$；如果$N_C= N_E(A')$，设置$P_R(A')=P$，$N_C=N_S(A)$，转到第四步。

如图 6-19-c 所示，如果从弧段 4 开始搜寻，找到节点 C 后，根据弧段的排序，下一条弧段是 2；然后找到节点 A，弧段 1，整个搜寻结束，建立多边形 I，其组成弧段为 4、2、1。

按照这种算法，生成多边形的弧段从多边形内部看，是逆时针排列的。如果节点弧段排序为顺时针，则算法中用$P_L(A)$代替$P_R(A)$，用$P_R(A)$代替$P_L(A)$，生成的多边形弧段是顺时针排列的。

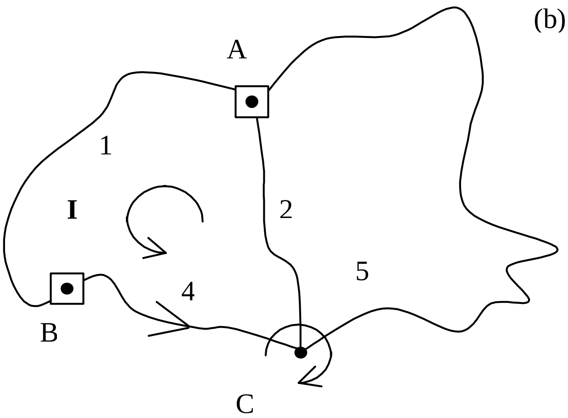


A	3	2	1
B	4	6	1
C	2	5	4
D	3	6	5

节点表，其中对于每个节点记录了顺时针方向排序的相连的弧段

(a)

(b)



(c)

图 6-19：多边形拓扑的建立过程

多边形拓扑的建立，要注意多边形带“岛”的情况，按照上述算法，对于带“岛”的多边形，或者称为环，其包含的弧段构成了多个闭合曲线，并且“岛”的弧段排序是顺时针的（图 6-20）（实际上，从环状多边形内部看，它仍然是逆时针的）。

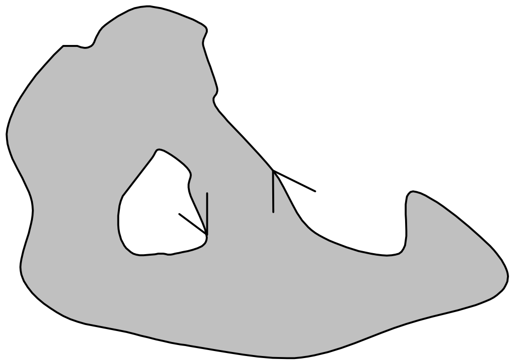


图 6-20：带“岛”的多边形建立拓扑的结果。

2）网络拓扑关系的建立

在输入道路、水系、管网、通信线路等信息时，为了进行流量以及连通性分析，需要确定线实体之间的连接关系。网络拓扑关系的建立包括确定节点与连接线之间的关系，这个工作可以由计算机自动完成，但是在一些情况中，如道路交通应用中，一些道路虽然在平面上相交，但是实际上并不连通，如立交桥，这是需要手工修改，将连通的节点删除（图 6-21）。

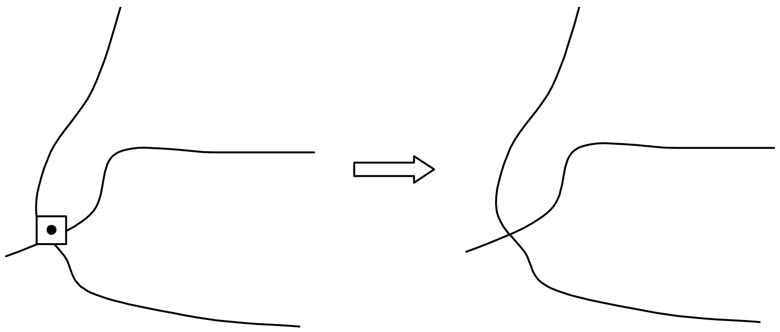


图 6-21：节点的编辑，将实际不连通的线路形成的节点删除

妈妈说："生活就像是一盒巧克力，你永远不知道下一块的味道。"

Forrest Gump

第七章 空间数据管理

导读：本章首先介绍空间数据库、与一般数据库的比较，以及空间数据库的存储方式。

然后介绍了 GIS 中两种重要的数据结构：栅格结构和矢量结构，以及其具体的存储方式，然后比较了两种结构的特点，并给出了其相互转换算法。

最后介绍了空间检索中常用的技术——空间索引，介绍了一些常用的空间索引方式，如 BSP 树、R 树、CELL 树等；以及空间数据的查询功能。

1．空间数据库

一个信息系统及其数据库的组成，决定于系统的应用目的、数据类型和系统的工作方式。关于地理信息系统的内容及其功能，以及地理信息系统的一个重要特点，或者说是与一般管理信息系统的区别，是数据具有空间分布的性质。对地理信息系统来讲，不仅数据本身具有空间属性，系统的分析和应用也无不与地理环境直接关联。系统的这一基本特征，深刻地影响着数据的结构、数据库的设计、分析算法和软件，以及系统的输入和输出。

1．1 地理信息系统与一般管理信息系统的比较

从数据源的角度来看，图形和图像数据是地理信息系统数据的一个主要来源，分析处理的结果也常用图形的方式来表示。而一般的管理信息系统，则多以统计数据、表格数据为主。这一点也使地理信息系统在硬件和软件上与一般的管理信息系统有所区别。

1．1．1 两者的区别

1）在硬件上，为了处理图形和图像数据，系统需要配置专门的输入和输出设备，如数字化仪、绘图机、图形图像的显示设备等；许多野外实地采集和台站的观测所得到的资源信息是模拟量形式，系统还需要配置模——数转换设备，这些设备往往超过中央处理机的价格，体积也比较大。

2）在软件上，则要求研制专门的图形和图像数据的分析算法和处理软件，这些算法和软件又直接和数据的结构及数据库的管理方法有关。

3）在信息处理的内容和采用目的方面，一般的管理信息系统，主要是查询检索和统计分析，处理的结果，大多是制成某种规定格式的表格数据，而地理信息系统，除了基本的信息检索和统计分析外，主要用于分析研究资源的合理开发利用，制定区域发展规划，地区的综合治理方案，对环境进行动态的监视和预测预报，为国民经济建设中的决策提供科学依据，为生产实践提供信息和指导。

由于地理信息系统是一个复杂的自然和社会的综合体，所以信息的处理必然是多因素的综合分析。系统分析是基本的方法，例如，研究某种地理信息系统中各组成部分间的相互关系，利用统计数据建立系统的数学模型，根据给定的目标函数，进行数学规划，寻求最优方

案，使该系统的经济效益为最佳；或者分析系统中各部分之间反馈联系，建立系统的结构模型，采用系统动力学的方法，进行动态分析，研究系统状态的变化和预测发展趋势等。计算机仿真是一种有效而经济的分析方法，便于分析各种因素的影响和进行方案的比较，在自然环境和社会经济的许多应用研究中常被采用。此外，地理信息系统还有分析量算的功能，如计算面积、长度、密度、分布特征等以及地理实体之间的关系运算。

1．1．2 两者共同之处

地理信息系统和一般的信息管理系统，也有许多共同之处。两者都是以计算机为核心的信息处理系统，都具有数据量大和数据之间关系复杂的特点，也都随着数据库技术的发展在不断的改进和完善。比较起来，商用的管理信息系统发展快，用户数量大，而且已有定型的软件产品可供选用，这也促进了软件系统的标准化。地理信息系统，由于上述一些特点，多是根据具体的应用要求专门设计，数据格式和组织管理方法各不相同。目前国外已有几百个空间数据处理系统和软件包，几乎没有两个系统是一样的，尽管大家都认为标准化是很重要的，也作了许多努力（例如建立计算机制图的标准和规范），但分析的算法和软件系统还谈不上标准化的问题。事实上，地理信息系统正作为一种空间信息的处理系统，成为一个单独的研究和发展领域。

1．2 空间数据库

1．2．1 数据库的概念

数据库就是为一定目的服务，以特定的数据存储的相关联的数据集合，它是数据管理的高级阶段，是从文件管理系统发展而来的。地理信息系统的数据库（简称空间数据库或地理数据库）是某一区域内关于一定地理要素特征的数据集合。为了直观地理解数据库，可以把数据库作如下比较：

表 7-1：数据库与图书馆比较

数据库	图书馆
数据	图书
数据模型	书卡编目
数据的物理组织	图书存放规则、书架
数据库管理系统	图书管理员
外存	书库
用户	读者
数据存取	图书阅览

1．2．2 空间数据库特点

空间数据库与一般数据库相比，具有以下特点：

1）数据量特别大，地理系统是一个复杂的综合体，要用数据来描述各种地理要素，尤其是要素的空间位置，其数据量往往很大。

2）不仅有地理要素的属性数据（与一般数据库中的数据性质相似），还有大量的空间数

据，即描述地理要素空间分布位置的数据，并且这两种数据之间具有不可分割的联系。

3）数据应用广泛，例如地理研究、环境保护、土地利用与规划、资源开发、生态环境、市政管理、道路建设等。

1．2．3 数据库管理系统

数据库是关于事物及其关系的信息组合，早期的数据库物体本身与其属性是分开存储的，只能满足简单的数据恢复和使用。数据定义使用特定的数据结构定义，利用文件形式存储，称之为文件处理系统。

文件处理系统是数据库管理最普遍的方法，但是有很多缺点：首先每个应用程序都必须直接访问所使用的数据文件，应用程序完全依赖于数据文件的存储结构，数据文件修改时应用程序也随之修改；另外的问题是数据文件的共享。由于若干用户或应用程序共享一个数据文件，要修改数据文件必须征得所有用户的认可。由于缺乏集中控制也会带来一系列数据库的安全问题。数据库的完整性是严格的，信息质量很差比没有信息更糟。

数据库管理系统（Database Management System，DBMS）是在文件处理系统的基础上进一步发展的系统。DBMS 在用户应用程序和数据文件之间起到了桥梁作用。DBMS 的最大优点是提供了两者之间的数据独立性，即应用程序访问数据文件时，不必知道数据文件的物理存储结构。当数据文件的存储结构改变时，不必改变应用程序。

1）采用标准DBMS存储空间数据的主要问题*

用标准的 DBMS 来存储空间数据，不如存储表格数据那样好，其主要问题包括：

（1．1）在 GIS 中，空间数据记录是变长的，因为需要存储的坐标点的数目是变化的，而一般数据库都只允许把记录的长度设定为固定长度。不仅如此，在存储和维护空间数据拓扑关系方面，DBMS 也存在着严重的缺陷。因而，一般要对标准的 DBMS 增加附加的软件功能。

（1．2）DBMS 一般都难以实现对空间数据的关联、连通、包含、叠加等基本操作。

（1．3）GIS 需要一些复杂的图形功能，一般的 DBMS 不能支持。

（1．4）地理信息是复杂的，单个地理实体的表达需要多个文件、多条记录、或许包括大地网、特征坐标、拓扑关系、空间特征量测值、属性数据的关键字以及非空间专题属性等，一般的 DBMS 也难以支持。

（1．5）具有高度内部联系的 GIS 数据记录需要更复杂的安全性维护系统，为了保证空间数据库的完整性，保护数据文件的完整性，保护系列必须与空间数据一起存储，否则一条记录的改变就会使其他数据文件产生错误。一般的 DBMS 都难以保证这些。

2）GIS 数据管理方法主要 4 种类型

（2．1）对不同的应用模型开发独立的数据管理服务，这是一种基于文件管理的处理方法。

（2．2）在商业化的 DBMS 基础上开发附加系统。开发一个附加软件用于存储和管理空间数据和空间分析，使用 DBMS 管理属性数据。

（2．3）使用现有的 DBMS，通常是以 DBMS 为核心，对系统的功能进行必要扩充，空间数据和属性数据在同一个 DBMS 管理之下。需要增加足够数量的软件和功能来提供空间功能和图形显示功能。

（2．4）重新设计一个具有空间数据和属性数据管理和分析功能的数据库系统。

---

* 在“地理信息系统软件工程技术”一章中“数据管理设计”中详细描述了各种不同的数据管理方案。

1．3 数据与文件组织

数据是现实世界中信息的载体，是信息的具体表达形式，为了表达有意义的信息内容，数据必须按照一定的方式进行组织和存储。

1．3．1 数据组织的分级

数据库中的数据组织一般可以分为四级：数据项、记录、文件和数据库。

1）数据项

数据项是可以定义数据的最小单位，也叫元素、基本项、字段等，数据项与现实世界实体的属性相对应，数据项有一定的取值范围，称为域，域以外的任何值对该数据项都是无意义的。每个数据项都有一个名称，称为数据项目。数据项的值可以是数值的、字母的、字母数字的、汉字的等形式。数据项的物理特点在于它具有确定的物理长度，可以作为整体看待。

2）记录

记录是由若干相关联的数据项组成，是处理和存储信息的基本单位，是关于一个实体的数据总和，构成该记录的数据项表示实体的若干属性。记录有“型”和“值”的区别，“型”是同类记录的框架，它定义记录；而“值”是记录反映实体的内容。为了唯一标识每个记录，就必须有记录标识符，也叫关键字。记录标识符一般由记录中的第一个数据项担任，唯一标识记录的关键字称主关键字，其它标识记录的关键字称为辅关键字。记录可以分为逻辑记录与物理记录，逻辑记录是文件中按信息在逻辑上的独立意义来划分的数据单位；而物理记录是单个输入输出命令进行数据存取的基本单元。物理记录和逻辑记录之间的对应关系有一个物理记录一对应一个逻辑记录；一个物理记录含有若干个逻辑记录；若干个物理记录存放一个逻辑记录。

3）文件

文件是一给定类型的（逻辑）记录的全部具体值的集合，文件用文件名称标识，文件根据记录的组织方式和存取方法可以分为：顺序文件、索引文件、直接文件和倒排文件等。

4）数据库

数据库是比文件更大的数据组织，数据库是具有特定联系的数据的集合，也可以看成是具有特定联系的多种类型的记录的集合。数据库的内部构造是文件的集合，这些文件之间存在某种联系，不能孤立存在。

1．3．2 数据间的逻辑联系

数据间的逻辑联系主要是指记录与记录之间的联系。记录是表示现实世界中的实体的。实体之间存在着一种或多种联系，这样的联系必然要反映到记录之间的联系上来。数据之间的逻辑联系主要有三种：一对一的联系；一对多的联系；多对多的联系。

1．3．3 常用数据文件

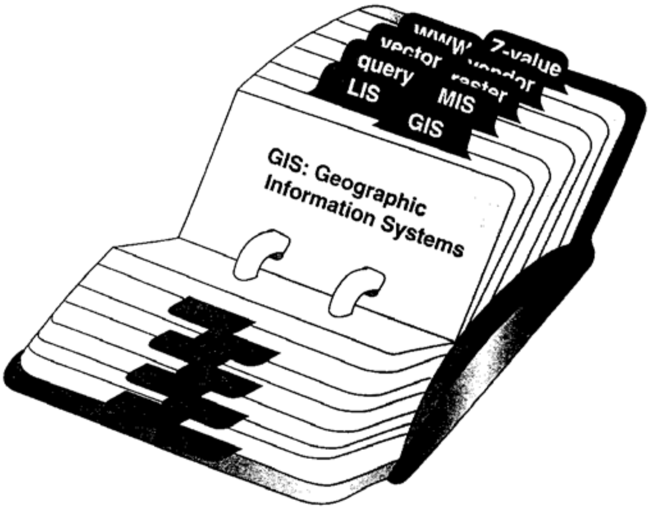


图 7-1：非顺序文件

文件组织是数据组织的一部分，数据组织既指数据在内存中的组织，又指数据在外存中的组织，而文件组织则主要指数据记录在外存设备上的组织，它由操作系统 OS 进行管理，具体讲在外存设备上如何安排数据和组织数据，以及实施对数据的访问方式等问题。操作系统实现的文件组织方式，可以分为顺序文件、索引文件、直接文件和倒排文件。

1）顺序文件

顺序文件（图 7-2）是最简单的文件组织形式，对记录按照主关键字的顺序进行组织。当主关键字是数字型时，以其数值的大小为序；若主关键字是文字型的，则以字母的排列为序。一切存于磁带上的记录，都只能是顺序的，而存于磁盘上的记录，既可以是顺序的，也可以是随机的。顺序文件的记录，逻辑上是按主关键字排序的，而在物理存储上可以有不同的方式，包括向量方式：被存储的文件按地址连续存放，物理结构与逻辑结构一致；链方式：文件不按地址连续存放，文件的逻辑顺序靠链来实现，文件中的每个记录中都含有一个指针，用以指明下一个记录的存放地址；块链方式：把文件分成若干数据块，块之间用指针连接，而块内则是连续存储。

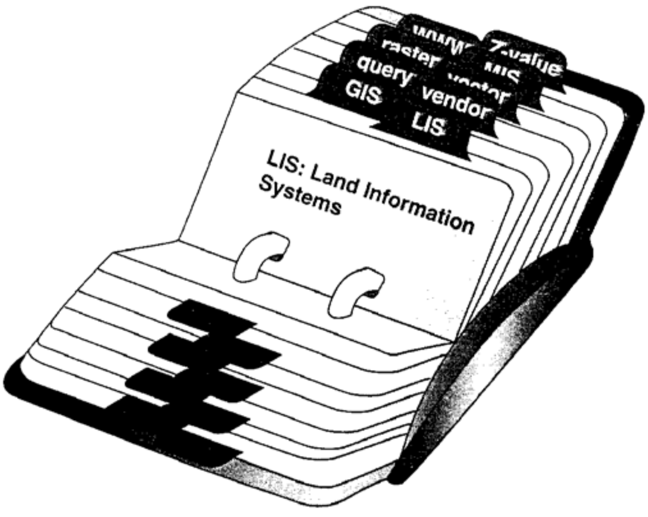


图 7-2：顺序文件

2）索引文件

索引文件除了存储记录本身（主文件）以外，还建立了若干索引表，这种带有索引表的文件叫索引文件。索引表中列出记录关键字和记录在文件中的位置（地址）。读取记录时，只要提供记录的关键字值，系统通过查找索引表获得记录的位置，然后取出该记录。索引表一般都是经过排序的，既可以是有顺序的，也可以是非顺序的，可以是单级索引，也可以是多级索引，多级索引可以提高查找速度，但占用的存储空间较大。

3）直接文件

直接文件又称随机文件，其存储是根据记录关键字的值，通过某种转换方法得到一个物理存储位置，然后把记录存储在该位置上。查找时，通过同样的转换方法，可以直接得到所需要的记录。

4）倒排文件

倒排文件是带有辅索引的文件，其中辅索引是按照一些辅关键字来组织索引的（注意：索引文件是按照记录的主关键字来构造索引的，也叫主索引）。倒排文件是一种多关键字的索引文件，其中的索引不能唯一标识记录，往往同一索引指向若干记录。因而，索引往往带有一个指针表，指向所有该索引标识的记录。通过辅索引不能直接读取记录，而要通过主关键字才能查到记录的位置。倒排文件的主要优点是在处理多索引检索时，可以在辅检索中先完成查询的‘交’、‘并’等逻辑运算，得到结果后再对记录进行存取，从而提高查找速度。

1．4 GIS 的内部数据结构——矢量结构和栅格结构

描述地理实体的数据本身的组织方法，称为内部数据结构。空间数据结构是指适合于计算机系统存储、管理和处理的地学图形的逻辑结构，是地理实体的空间排列方式和相互关系的抽象描述。它是对数据的一种理解和解释，不说明数据结构的数据是毫无用处的，不仅用户无法理解，计算机程序也不能正确的处理。对同样的一组数据，按不同的数据结构去处理，得到的可能是截然不同的内容。空间数据结构是地理信息系统沟通信息的桥梁，只有充分理解地理信息系统所采用的特定数据结构，才能正确地使用系统。

内部数据结构基本上可分为两大类：矢量结构和栅格结构（也可以称为矢量模型和栅格模型）（图 7-3）。两类结构都可用来描述地理实体的点、线、面三种基本类型。

空间数据编码是空间数据结构的实现，即将根据地理信息系统的目的和任务所搜集的、经过审核了的地形图、专题地图和遥感影像等资料按特定的数据结构转换为适合于计算机存储和处理的数据的过程。由于地理信息系统数据量极大，一般采用压缩数据的编码方式以减少数据冗余。

在地理信息系统的空间数据结构中，栅格结构的编码方式主要有直接栅格编码、链码、游程长度编码、块码、四叉树码等；矢量结构主要有坐标序列编码、树状索引编码和二元拓扑编码等编码方法。

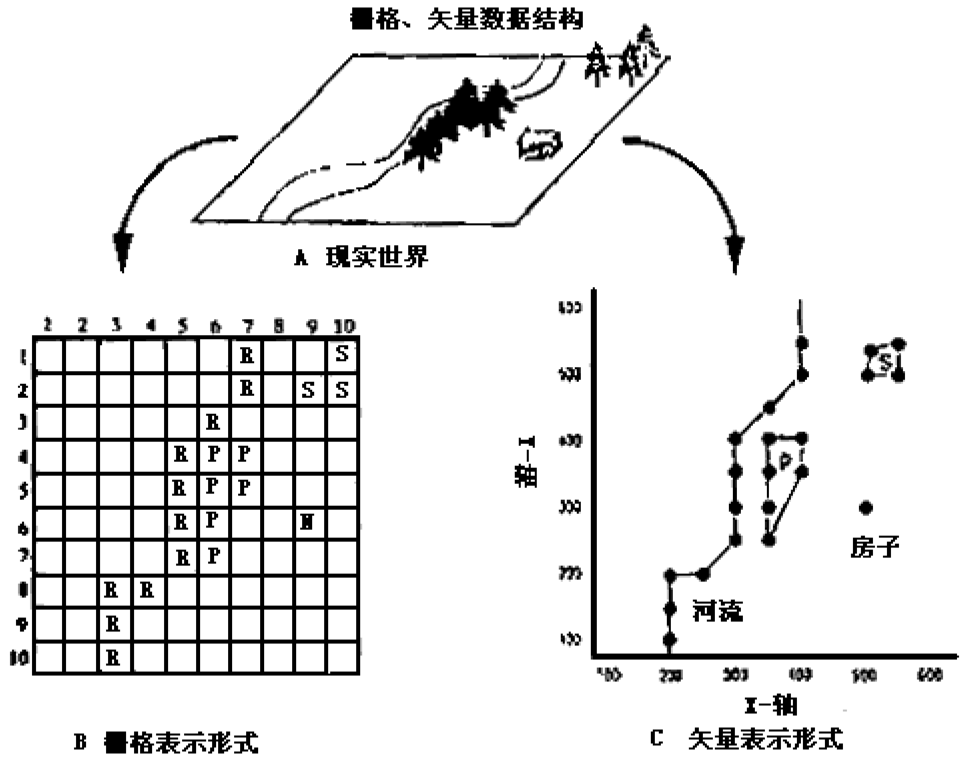


图 7-3：矢量结构和栅格结构

1．4．1 矢量模型

在矢量模型中，现实世界的要素位置和范围可以采用点、线或面表达，与它们在地图上表示相似，每一个实体的位置是用它们在坐标参考系统中的空间位置（坐标）定义。地图空间中的每一位置都有唯一的坐标值。点、线和多边形用于表达不规则的地理实体在现实世界的状态（多边形是由若干直线围成的封闭区域的边界）。一条线可能表达一条道路，一个多边形可能表达一块林地等。矢量模型中的空间实体与要表达的现实世界中的空间实体具有一定的对应关系。

1．4．2 栅格模型

在栅格模型中，空间被规则地划分为栅格（通常为正方形）。地理实体的位置和状态是用它们占据的栅格的行、列来定义的。每个栅格的大小代表了定义的空间分辨率。由于位置是由栅格行列号定义的，所以特定的位置由距它最近的栅格记录决定。例如，某个区域被划分成 10*10 个栅格，那么仅能记录位于这 10*10 个栅格附近的物体的位置。栅格的值表达了这个位置上物体的类型或状态。采用栅格方法，空间被划分成大量规则格网，而且每个栅格取值可能不一样。空间单元是栅格，每一个栅格对应于一个特定的空间位置，如地表的一个区域，栅格的值表达了这个位置的状态。

与矢量模型不一样，栅格模型最小单元与它表达的真实世界空间实体没有直接的对应关系。栅格数据模型中的空间实体单元不是通常概念上理解的物体，它们只是彼此分离的栅格。例如，道路作为明晰的栅格是不存在的，栅格的值才表达了路是一个实体。道路是被具有道路属性值的一组栅格表达的，这条路不可能通过某一栅格实体被识别出来。在这两种数据结构中，空间信息都是使用统一的单位表达。在栅格方法中，统一的单位是栅格（栅格是不可

再分的，其属性用于表达对应位置物体的性质），表达一个区域所用栅格的数量很大，但其栅格单元的大小一样。栅格数据文件包含有上百万个栅格，每个栅格的位置都被严格定义。在矢量方法中，统一的单元是点、线和多边形，与栅格方法相比，在数量上所用的表达单元较少，但大小可变。在矢量文件中，元素的个数或许数千个，但毕竟没有栅格数据那么多。同一类型的矢量单元的位置是用连续坐标值定义。矢量数据提供的坐标位置比栅格数据用行、列号所表达位置更精确。这两种方法各有优缺点，究竟采用何种数据结构，取决于利用数据的目的。有些地理现象用栅格数据表达更合适；有些地理现象则用矢量数据更有利，以便表达它们之间的空间关系。

2．栅格数据结构及其编码

2．1 栅格数据结构

2．1．1 定义

栅格结构是最简单最直接的空间数据结构，是指将地球表面划分为大小均匀紧密相邻的网格阵列，每个网格作为一个象元或象素由行、列定义，并包含一个代码表示该象素的属性类型或量值，或仅仅包括指向其属性记录的指针。因此，栅格结构是以规则的阵列来表示空间地物或现象分布的数据组织，组织中的每个数据表示地物或现象的非几何属性特征。如图 7-4 所示，在栅格结构中，点用一个栅格单元表示；线状地物沿线走向的一组相邻栅格单元表示，每个栅格单元最多只有两个相邻单元在线上；面或区域用记有区域属性的相邻栅格单元的集合表示，每个栅格单元可有多于两个的相邻单元同属一个区域。遥感影像属于典型的栅格结构，每个象元的数字表示影像的灰度等级。

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	2	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

（a）点

0	0	0	0	0	0	0	0
0	0	0	6	0	0	0	0
0	6	6	0	6	0	0	0
0	0	0	0	0	6	0	0
0	0	0	0	0	6	0	0
0	0	0	0	0	6	0	0
0	0	0	0	0	0	6	0
0	0	0	0	0	0	0	0

（b）线

0	4	4	7	7	7	7	7
4	4	4	4	4	7	7	7
4	4	4	4	8	8	7	7
0	0	4	8	8	8	7	7
0	0	8	8	8	8	7	8
0	0	0	8	8	8	8	8
0	0	0	0	8	8	8	8
0	0	0	0	0	8	8	8

（c）面

图 7-4：点、线、区域的格网

2．1．2 特点

栅格结构的显著特点是：属性明显，定位隐含，即数据直接记录属性的指针或属性本身，而所在位置则根据行列号转换为相应的坐标，也就是说定位是根据数据在数据集中的位置得到的。如图 7-4-（a）所示，数据 2 表示属性或编码为 2 的一个点，其位置由其所在的第 3 行、第 4 列交叉得到的。由于栅格结构是按一定的规则排列的，所表示的实体的位置很容易

隐含在格网文件的存储结构中，在后面讲述栅格结构编码时可以看到，每个存储单元的行列位置可以方便地根据其在文件中的记录位置得到，且行列坐标可以很容易地转为其他坐标系下的坐标。在格网文件中每个代码本身明确地代表了实体的属性或属性的编码，如果为属性的编码，则该编码可作为指向实体属性表的指针。图 7-4-（a）表示了代码为 2 的点实体，图 7-4-（b）表示了一条代码为 6 的线实体，而图 7-4-（c）则表示了三个面实体或称为区域实体，代码分别为 4、7 和 8。由于栅格行列阵列容易为计算机存储、操作和显示，因此这种结构容易实现，算法简单，且易于扩充、修改，也很直观，特别是易于同遥感影像的结合处理，给地理空间数据处理带来了极大的方便。

栅格结构表示的地表是不连续的，是量化和近似离散的数据。在栅格结构中，地表被分成相互邻接、规则排列的矩形方块（特殊的情况下也可以是三角形或菱形、六边形等），每个地块与一个栅格单元相对应。栅格数据的比例尺就是栅格大小与地表相应单元大小之比。在许多栅格数据处理时，常假设栅格所表示的量化表面是连续的，以便使用某些连续函数。由于栅格结构对地表的量化，在计算面积、长度、距离、形状等空间指标时，若栅格尺寸较大，则造成较大的误差，由于在一个栅格的地表范围内，可能存在多于一种的地物，而表示在相应的栅格结构中常常是一个代码。也类似于遥感影像的混合象元问题，如 Landsat 的 MSS 卫星影像单个象元对应地表 79 米*79 米的矩形区域，影像上记录的光谱数据是每个象元所对应的地表区域内所有地物类型的光谱辐射的总和效果。因而，这种误差不仅有形态上的畸形，还可能包括属性方面的偏差。

2．2 决定栅格单元代码的方式

在决定栅格代码时尽量保持地表的真实性，保证最大的信息容量。图 7-5 所示的一块矩形地表区域，内部含有 A、B、C 三种地物类型，O 点为中心点，将这个矩形区域近似地表示为栅格结构中的一个栅格单元时，可根据需要，采取如下的方式之一来决定栅格单元的代码。

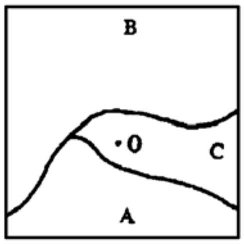


图 7-5：栅格单元代码的确定

2．2．1 中心点法

用处于栅格中心处的地物类型或现象特性决定栅格代码，在图 7-5 所示的矩形区域中，中心点 O 落在代码为 C 的地物范围内，按中心点法的规则，该矩形区域相应的栅格单元代码为 C，中心点法常用于具有连续分布特性的地理要素，如降雨量分布、人口密度图等。

2．2．2 面积占优法

以占矩形区域面积最大的地物类型或现象特性决定栅格单元的代码，在图 7-5 所示的例

子中，显见 B 类地物所占面积最大，故相应栅格代码定为 B。面积占优法常用于分类较细，地物类别斑块较小的情况。

2．2．3 重要性法

根据栅格内不同地物的重要性，选取最重要的地物类型决定相应的栅格单元代码，假设图 7-5 中 A 类最重要的地物类型，即 A 比 B 和 C 类更为重要，则栅格单元的代码应为 A。重要性法常用于具有特殊意义而面积较小的地理要素，特别是点、线状地理要素，如城镇、交通枢纽、交通线、河流水系等，在栅格中代码应尽量表示这些重要地物。

2．2．4 百分比法

根据矩形区域内各地理要素所占面积的百分比数确定栅格单元的代码，如可记面积最大的两类 BA，也可以根据 B 类和 A 类所占面积百分比数在代码中加入数字。

2．3 编码方法

2．3．1 直接栅格编码

这是最简单直观而又非常重要的一种栅格结构编码方法，通常称这种编码的图像文件为网格文件或栅格文件，栅格结构不论采用何种压缩编码方法，其逻辑原型都是直接编码网格文件。直接编码就是将栅格数据看作一个数据矩阵，逐行（或逐列）逐个记录代码，可以每行都从左到右逐个象元记录，也可以奇数行地从左到右而偶数行地从右向左记录，为了特定目的还可采用其他特殊的顺序（图 7-6）。

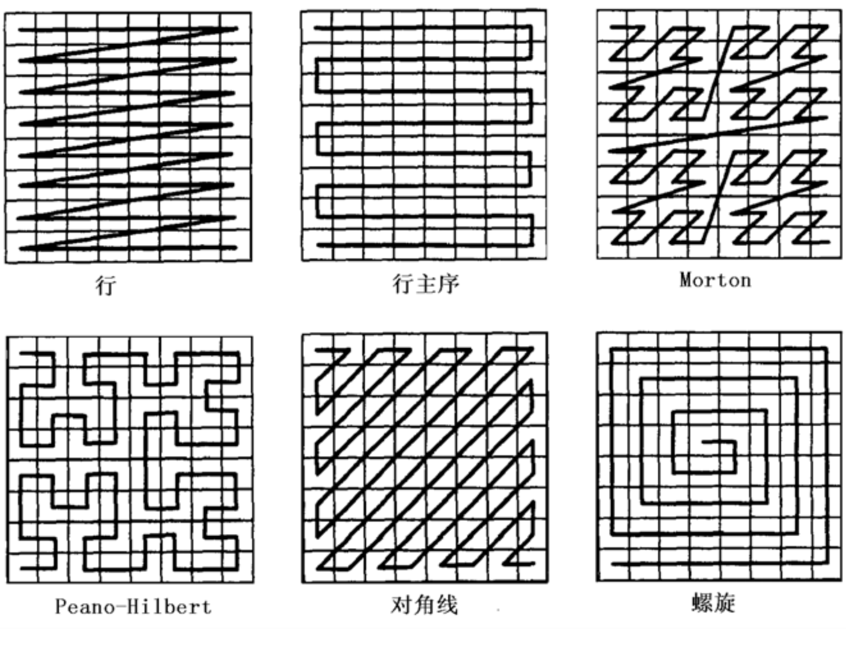


图 7-6：一些常用的栅格排列顺序

2．3．2 压缩编码方法

目前有一系列栅格数据压缩编码方法，如键码、游程长度编码、块码和四叉树编码等。其目的，就是用尽可能少的数据量记录尽可能多的信息，其类型又有信息无损编码和信息有损编码之分。信息无损编码是指编码过程中没有任何信息损失，通过解码操作可以完全恢复原来的信息，信息有损编码是指为了提高编码效率，最大限度地压缩数据，在压缩过程中损失一部分相对不太重要的信息，解码时这部分难以恢复。在地理信息系统中多采用信息无损编码，而对原始遥感影像进行压缩编码时，有时也采取有损压缩编码方法。

1）链码（Chain Codes）

链码又称为弗里曼链码[Freeman]或边界链码，链码可以有效地压缩栅格数据，而且对于估算面积、长度、转折方向的凹凸度等运算十分方便，比较适合于存储图形数据。缺点是对边界进行合并和插入等修改编辑工作比较困难，对局部的修改将改变整体结构，效率较低，而且由于链码以每个区域为单位存储边界，相邻区域的边界将被重复存储而产生冗余。

2）游程长度编码（Run-Length Codes）

游程长度编码是栅格数据压缩的重要编码方法，它的基本思路是：对于一幅栅格图像，常常有行（或列）方向上相邻的若干点具有相同的属性代码，因而可采取某种方法压缩那些重复的记录内容。其方法有两种方案：一种编码方案是，只在各行（或列）数据的代码发生变化时依次记录该代码以及相同的代码重复的个数，从而实现数据的压缩。例如对图 7-4（c）所示栅格数据，可沿行方向进行如下游程长度编码：

（0，1），（4，2），（7，5）；（4，5），（7，3）；（4，4），（8，2），（7，2）；（0，2），（4，1），（8，3），（7，2）；（0，2），（8，4），（7，1），（8，1）；(0，3)，(8，5)；(0，4)，(8，4)；(0，5)，(8，3)。

只用了 44 个整数就可以表示，而在前述的直接编码中却须要 64 个整数表示，可见游程长度编码压缩数据是十分有效又简便的。事实上，压缩比的大小是与图的复杂程度成反比的，在变化多的部分，游程数就多，变化少的部分游程数就少，图件越简单，压缩效率就越高。另一种游程长度编码方案就是逐个记录各行（或列）代码发生变化的位置和相应代码，如对图 7-4（c）所示栅格数据的另一种游程长度编码如下（沿列方向）：

（1，0），（2，4），（4，0），（1，4），（4，0）；（1，4），（5，8），（6，0）；（1，7），（2，4），（4，8），（7，0）；（1，7），（2，4），（3，8），（8，0）；（1，7），（3，8）；（1，7），（6，8）；（1，7），（5，8）。

游程长度编码在栅格压缩时，数据量没有明显增加，压缩效率较高，且易于检索，叠加合并等操作，运算简单，适用于机器存储容量小，数据需大量压缩，而又要避免复杂的编码解码运算增加处理和操作时间的情况。

3）块码

块码是游程长度编码扩展到二维的情况，采用方形区域作为记录单元，每个记录单元包括相邻的若干栅格，数据结构由初始位置（行、列号）和半径，再加上记录单位的代码组成。对图 7-4（c）所示图像的块码编码如下：

(1,1,1,0),(1,2,2,4),(1,4,1,7),(1,5,1,7),
(1,6,2,7),(1,8,1,7),(2,1,1,4),(2,4,1,4),
(2,5,1,4),(2,8,1,7),(3,1,1,4),(3,2,1,4),
(3,3,1,4),(3,4,1,4),(3,5,2,8),(3,7,2,7),
(4,1,2,0),(4,3,1,4),(4,4,1,8),(5,3,1,8),

(5,4,2,8),(5,6,1,8),(5,7,1,7),(5,8,1,8),

(6,1,3,0),(6,6,3,8),(7,4,1,0),(7,5,1,8),

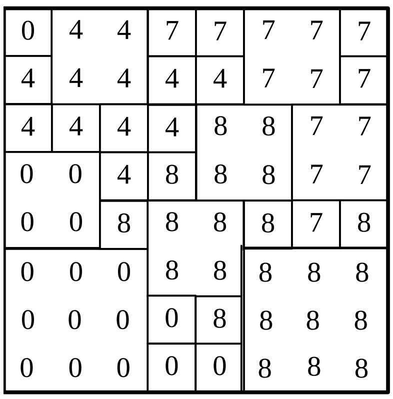
(8,4,1,0),(8,5,1,0)。

该例中块码用了 120 个整数，比直接编码还多，这是因为例中为描述方便，栅格划分很粗糙，在实际应用中，栅格划分细，数据冗余多的多，才能显出压缩编码的效果，而且还可以作一些技术处理，如行号可以通过行间标记而省去记录，行号和半径等也不必用双字节整数来记录，可进一步减少数据冗余。

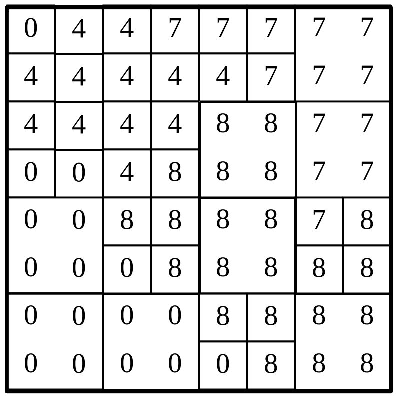
块码具有可变的分辨率，即当代码变化小时图块大，就是说在区域图斑内部分辨率低；反之，分辨率高以小块记录区域边界地段，以此达到压缩的目的。因此块码与游程长度编码相似，随着图形复杂程度的提高而降低效率，就是说图斑越大，压缩比越高；图斑越碎，压缩比越低。块码在合并、插入、检查延伸性、计算面积等操作时有明显的优越性。然而在某些操作时，则必须把游程长度编码和块码解码，转换为基本栅格结构进行。

4）四叉树

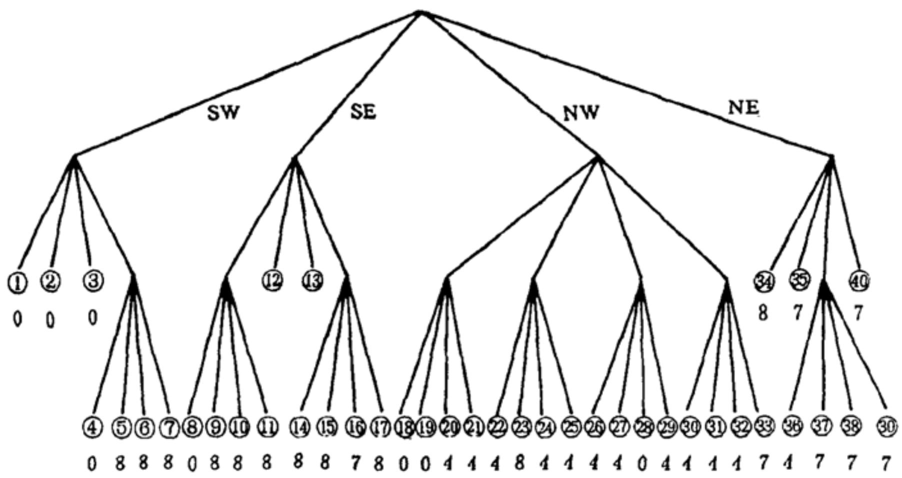
四叉树又称四元树或四分树，是最有效的栅格数据压缩编码方法之一，绝大部分图形操作和运算都可以直接在四叉树结构上实现，因此四叉树编码既压缩了数据量，又可大大提高图形操作的效率。四叉树将整个图像区逐步分解为一系列被单一类型区域内含的方形区域，最小的方形区域为一个栅格象元，分割的原则是，将图像区域划分为四个大小相同的象限，而每个象限又可根据一定规则判断是否继续等分为次一层的四个象限，其终止判据是，不管是哪一层上的象限，只要划分到仅代表一种地物或符合既定要求的少数几种地物时，则不再继续划分，否则一直划分到单个栅格象元为止。四叉树通过树状结构记录这种划分，并通过这种四叉树状结构实现查询、修改、量算等操作。图 7-7（b）为图 7-4（c）图形的四叉树分解，各子象限尺度大小不完全一样，但都是同代码栅格单元，其四叉树如图 7-7-（c）所示。



（a）块码分割



（b）四叉树分割



（c）b 的四叉树编码

图 7-7：四叉树编码

其中最上面的那个结点叫做根结点，它对应整个图形。总共有 4 层结点，每个结点对应一个象限，如 2 层 4 个结点分别对应于整个图形的四个象限，排列次序依次为南西（SW）、南东（SE）、北西（NW）和北东（NE），不能再分的结点称为终止结点（又称叶子结点），可能落在不同的层上，该结点代表的子象限具有单一的代码，所有终止结点所代表的方形区域覆盖了整个图形。从上到下，从左到右为叶子结点编号如图 7-7(c)所示，共有 40 个叶子结点，也就是原图被划分为 40 个大小不等的方形子区，图 7-7(c)的最下面的一排数字表示各子区的代码。

由上面图形的四叉树分解可见，四叉树中象限的尺寸是大小不一的，位于较高层次的象限较大，深度小即分解次数少，而低层次上的象限较小，深度大即分解次数多，这反映了图上某些位置单一地物分布较广而另一些位置上的地物比较复杂，变化较大。正是由于四叉树编码能够自动地依照图形变化而调整象限尺寸，因此它具有极高的压缩效率。

采用四叉树编码时，为了保证四叉树分解能不断地进行下去，要求图像必须为 $2^n \times 2^n$ 的栅格阵列，n为极限分割数，n+1 为四叉树的最大高度或最大层数，图 7-4（c）为 $2^3 \times 2^3$ 的栅格，因此最多划分三次，最大层数为 4，对于非标准尺寸的图像需首先通过增加背景的方法将图像扩充为 $2^n \times 2^n$ 的图像。

为了使计算机既能以最小的冗余存储图像对应的四叉树，又能方便地完成各种图形图像操作，专家们已提出了多种编码方式，下面介绍美国马里兰大学地理信息系统中采用的编码方式，该方法记录了终止结点（或叶子结点）的地址和值，值就是子区的代码，其中地址包括两个部分，共 32 位（二进制），最右边 4 位记录该叶子结点的深度，即处于四叉树的第几层上，有了深度可以推知子区的大小；地址由从根结点到该叶子结点的路径表示，0，1，2，3 分别表示 SW、SE、NW、NE，从右边第 5 位开始 2n 字节记录这些方向。如图 7-7-（c）表示的第六个结点深度为 3，第一层处于 SW 象限，记为 0；第二层处于 NE 象限，记为 3，第三层处于 NW 象限，记为 2，表示为二进制为：

0000… 000(22 位)；001110（6 位）；0011（4 位）

每层象限位置由两位二进制数表示，共 6 位，十进制整数为 227。这样，记录了各个叶子的地址，再记上相应代码值，就记录了这个图像，并可在此编码基础上进行多种图像操作。

事实上，叶结点的地址可以直接由子区左下角的行列坐标，按二进制按位交错得到。如对于 6 号叶子结点，在以图像左下角为原点的行列坐标系中，其左下角行、列坐标为（3，2），表示为二进制分别为 011 和 010，按位交错就是 001110，正是 6 号地块。

对于只有点状地物或只有线状地物的图件，为了提高效率，设计了略有不同的划分终止

条件和记录方法，称为点四叉树和线四叉树。点四叉树对子象限的划分直到每个子象限不含有点或只含有一个点为止，叶子的值则记录是否有点和点在子象限的位置；线四叉树划分子象限直到子象限不含线段或只含有单个线段，对线的结点则划分到单个象素，其叶子值记录更为复杂。

四叉树编码具有可变的分辨率，并且有区域性质，压缩数据灵活，许多运算可以在编码数据上直接实现，大大地提高了运算效率，是优秀的栅格压缩编码之一。

一般说来，对数据的压缩是以增加运算时间为代价的。在这里时间与空间是一对矛盾，为了更有效地利用空间资源，减少数据冗余，不得不花费更多的运算时间进行编码，好的压缩编码方法就是要在尽可能减少运算时间的基础上达到最大的数据压缩效率，并且是算法适应性强，易于实现。链码的压缩效率较高，已经近矢量结构，对边界的运算比较方便，但不具有区域的性质，区域运算困难；游程长度编码既可以在很大程度上压缩数据，又最大限度地保留了原始栅格结构，编码解码十分容易；块码和四叉树码具有区域性质，又具有可变的分辨率，有较高的压缩效率，四叉树编码可以直接进行大量图形图像运算，效率较高，是很有前途的方法。在此基础上已经开始发展了用于三维数据的八叉树编码等。

3．矢量数据结构及其编码

3．1 矢量数据结构

3．1．1 定义

地理信息系统中另一种最常见的图形数据结构为矢量结构，即通过记录坐标的方式尽可能精确地表示点、线、多边形等地理实体，坐标空间设为连续，允许任意位置、长度和面积的精确定义，事实上，其精度仅受数字化设备的精度和数值记录字长的限制，在一般情况下，比栅格结构精度高得多。

对于点实体，矢量结构中只记录其在特定坐标系下的坐标和属性代码；对于线实体，在数字化时即进行量化，就是用一系列足够短的直线首尾相接表示一条曲线，当曲线被分割成多而短的线段后，这些小线段可以近似地看成直线段，而这条曲线也可以足够精确地由这些小直线段序列表示，矢量结构中只记录这些小线段的端点坐标，将曲线表示为一个坐标序列，坐标之间认为是以直线段相连，在一定精度范围内可以逼真地表示各种形状的线状地物；“多边形”在地理信息系统中是指一个任意形状、边界完全闭合的空间区域。其边界将整个空间划分为两个部分：包含无穷远点的部分称为外部，另一部分称为多边形内部。把这样的闭合区域称为多边形是由于区域的边界线同前面介绍的线实体一样，可以被看作是由一系列多而短的直线段组成，每个小线段作为这个区域的一条边，因此这种区域就可以看作是由这些边组成的多边形了。

跟踪式数字化仪对地图数字化产生矢量结构的数字地图，适合于矢量绘图仪绘出。矢量结构允许最复杂的数据以最小的数据冗余进行存储，相对栅格结构来说，数据精度高，所占空间小，是高效的空间数据结构。

3．1．2 特点

矢量结构的特点是：定位明显、属性隐含，其定位是根据坐标直接存储的，而属性则一般存于文件头或数据结构中某些特定的位置上，这种特点使得其图形运算的算法总体上比栅格数据结构复杂的多，有些甚至难以实现，当然有些地方也有所便利和独到之处，在计算长度、面积、形状和图形编辑、几何变换操作中，矢量结构有很高的效率和精度，而在叠加运算、邻域搜索等操作时则比较困难。

3．2 编码方法

3．2．1 点实体

对于点实体和线实体的矢量编码比较直接，只要能将空间信息和属性信息记录完全就可以了。点是空间上不能再分的地理实体，可以是具体的或抽象的，如地物点、文本位置点或线段网络的结点等，由一对 x、y 坐标表示。图 7-8-a 表示了点的矢量编码的基本内容。

3．2．2 线实体

线实体主要用来表示线状地物（如公路、水系、山脊线等）符号线和多边形边界，有时也称为“弧”、“链”、“串”等，其矢量编码一般包括以下内容，图 7-8-b 为线实体矢量编码的基本内容。

其中唯一标识码是系统排列序号；线标识码可以标识线的类型；起始点和终止点号可直接用坐标表示；显示信息是显示时的文本或符号等；与线相联系的非几何属性可以直接存储于线文件中，也可单独存储，而由标识码联接查找。

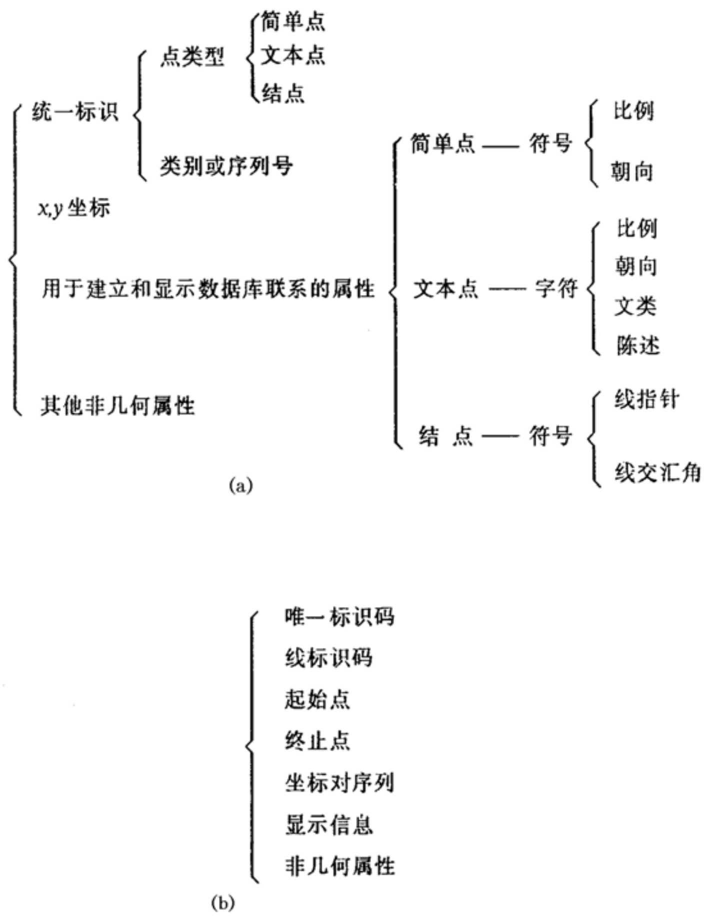


图 7-8：（a）点实体的编码，（b）线实体的编码

3．2．3 多边形

多边形数据是描述地理信息的最重要的一类数据。在区域实体中，具有名称属性和分类属性的，多用多边形表示，如行政区、土地类型、植被分布等；具有标量属性的，有时也用等值线描述（如地形、降雨量等)。

多边形矢量编码不但要表示位置和属性，更为重要的是要能表达区域的拓扑性质，如形状、邻域和层次等，以便使这些基本的空间单元可以作为专题图资料进行显示和操作，由于要表达的信息十分丰富，基于多边形的运算多而复杂，因此多边形矢量编码比点和线实体的矢量编码要复杂得多，也更为重要。

多边形矢量编码除有存储效率的要求外，一般还要求所表示的各多边形有各自独立的形状，可以计算各自的周长和面积等几何指标；各多边形拓扑关系的记录方式要一致，以便进行空间分析；要明确表示区域的层次，如岛-湖-岛的关系等。因此，它与机助制图系统仅为显示和制图目的而设计的编码有很大不同。

1）坐标序列法（Spaghetti 方式）

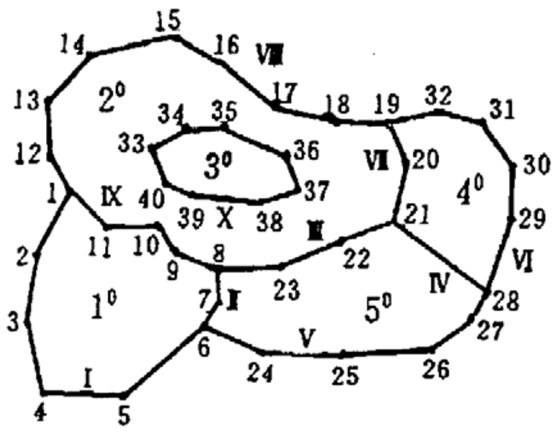


图 7-9：坐标序列法表示的多边形

由多边形边界的 x、y 坐标对集合及说明信息组成，是最简单的一种多边形矢量编码，如图 7-9 记为以下坐标文件：

1^0：x_1,y_1；x_2,y_2；x_3,y_3；x_4,y_4；x_5,y_5；x_6,y_6；x_7,y_7；x_8,y_8；x_9,y_9；x_{10},y_{10}；x_{11},y_{11}；

2^0：x_1,y_1；x_{12},y_{12}；x_{13},y_{13}；x_{14},y_{14}；x_{15},y_{15}；x_{16},y_{16}；x_{17},y_{17}；x_{18},y_{18}；x_{19},y_{19}；x_{20},y_{20}；x_{21},y_{21}；x_{22},y_{22}；x_{23},y_{23}；x_8,y_8；x_9,y_9；x_{10},y_{10}；x_{11},y_{11}；

3^0：x_{33},y_{33}；x_{34},y_{34}；x_{35},y_{35}；x_{36},y_{36}；x_{37},y_{37}；x_{38},y_{38}；x_{39},y_{39}；x_{40},y_{40}；

4^0：x_{19},y_{19}；x_{20},y_{20}；x_{21},y_{21}；x_{28},y_{28}；x_{29},y_{29}；x_{30},y_{30}；x_{31},y_{31}；x_{32},y_{32}；

5^0：x_{21},y_{21}；x_{22},y_{22}；x_{23},y_{23}；x_8,y_8；x_7,y_7；x_6,y_6；x_{24},y_{24}；x_{25},y_{25}；x_{26},y_{26}；x_{27},y_{27}；x_{28},y_{28}；

坐标序列法文件结构简单，易于实现以多边形为单位的运算和显示。这种方法的缺点是：

（1．1）多边形之间的公共边界被数字化和存储两次，由此产生冗余和碎屑多边形；

（1．2）每个多边形自成体系而缺少邻域信息，难以进行邻域处理，如消除某两个多边形之间的共同边界；

（1．3）岛只作为一个单个的图形建造，没有与外包多边形的联系；

（1．4）不易检查拓扑错误。这种方法可用于简单的粗精度制图系统中。

2）树状索引编码法

该法采用树状索引以减少数据冗余并间接增加邻域信息，方法是对所有边界点进行数字化，将坐标对以顺序方式存储，由点索引与边界线号相联系，以线索引与各多边形相联系，形成树状索引结构。

图 7-10 和图 7-11 分别为图 7-9 的多边形文件和线文件树状索引示意图。其文件结构如下：

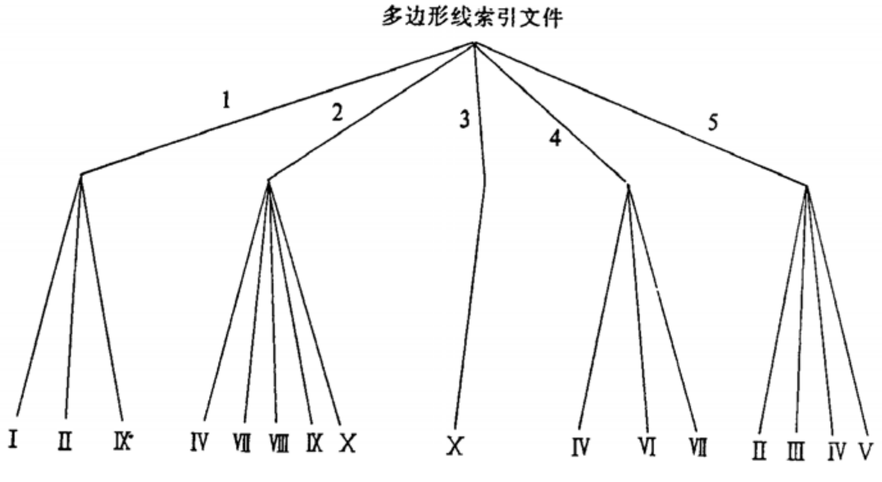


图 7-10：线与多边形之间的树状索引

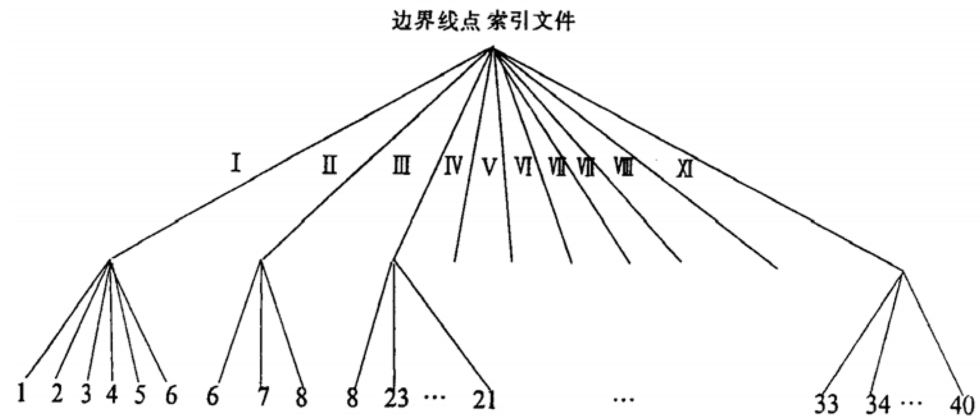


图 7-11：点与边界线之间的树状索引

采用上述的树状结构，图 7-9 的多边形数据记录如下：

1）点文件：

点号	坐标
1	x_1,y_1
2	x_2,y_2
…	…
40	x_{40},y_{40}

2）线文件

线号	起点	终点	点号
I	1	6	1,2,3,4,5,6
II	6	8	6,7,8
…	…	…	…
X	33	33	33,34,35,36,37,38,39,40,33

3）多边形文件

多边形编号	多边形边界
1^0	I,II,IX
2^0	III,VII,VIII,IX,X
3^0	X
4^0	IV,VI,VII
5^0	II,III,IV,V

树状索引编码消除了相邻多边形边界的数据冗余和不一致的问题，在简化过于复杂的边界线或合并相邻多边形时可不必改造索引表，邻域信息和岛状信息可以通过对多边形文件的线索引处理得到，但是比较繁琐，因而给相邻函数运算，消除无用边，处理岛状信息以及检查拓扑关系带来一定的困难，而且两个编码表都需要以人工方式建立，工作量大且容易出错。

3）拓扑结构编码法

要彻底解决邻域和岛状信息处理问题必须建立一个完整的拓扑关系结构，这种结构应包括以下内容：唯一标识，多边形标识，外包多边形指针，邻接多边形指针，边界链接，范围（最大和最小 x、y 坐标值）。采用拓扑结构编码可以较好地解决空间关系查询等问题，但增加了算法的复杂性和数据库的大小。

矢量编码最重要的是信息的完整性和运算的灵活性，这是由矢量结构自身的特点所决定的，目前并无统一的最佳的矢量结构编码方法，在具体工作中应根据数据的特点和任务的要求而灵活设计。

DIME(双重独立坐标地图编码，Dual Independent Map Encoding)编码系统

DIME 是美国人口调查局在人口调查的基础上发展起来的，它通过有向编码建立了多边形、边界、节点之间的拓扑关系，DIME 编码成为其它拓扑编码结构的基础。

4．矢栅结构的比较及转换算法

4．1 栅格结构与矢量结构的比较

栅格结构与矢量结构似乎是两种截然不同的空间数据结构，栅格结构“属性明显、位置隐含”，而矢量结构“位置明显、属性隐含”，栅格数据操作总的来说比较容易实现，尤其是作为斑块图件的表示更易于为人们接受；而矢量数据操作则比较复杂，许多分析操作（如两张地图的覆盖操作，点或线状地物的邻域搜索等）用矢量结构实现十分困难，矢量结构表达线状地物是比较直观的，而面状地物则是通过对边界的描述而表达。无论哪种结构，数据精度和数据量都是一对矛盾，要提高精度，栅格结构需要更多的栅格单元，而矢量结构则需记录更多的线段结点。一般来说，栅格结构只是矢量结构在某种程度上的一种近似，如果要使栅格结构描述的图件取得与矢量结构同样的精度，甚至仅仅在量值上接近，则数据也要比后者大得多。

栅格结构在某些操作上比矢量结构更有效更易于实现，如按空间坐标位置的搜索，对于栅格结构是极为方便的，而对矢量结构则搜索时间要长得多；在给定区域内的统计指标运算，包括计算多边形形状、面积、线密度、点密度，栅格结构可以很快算得结果，而采用矢量结构则由于所在区域边界限制条件难以提取而降低效率，对于给定范围的开窗、缩放栅格结构也比矢量结构优越；另一方面，矢量结构用于拓扑关系的搜索则更为高效，即诸如计算多边形形状搜索邻域、层次信息等；对于网络信息只有矢量结构才能完全描述；矢量结构在计算精度与数据量方面的优势也是矢量结构比栅格结构受到欢迎的原因之一，对图 7-10 而言，假设坐标精度要求为万分之一，即 5 位数字，采用矢量结构需记录 40 个结点，每个结点用两个双字节整数记录x、y坐标，加上对其他说明信息的描述，200 个字节足够了，而若用基本栅格记录，则需 $10000*10000=10^8$个字节，即使采用单字节记录栅格代码（不超过 255），也需约五百万个字节，当然实际图形的矢量结构记录采点一般要比图 7-10 密得多，但数据量仍大大少于栅格结构的数据量。

栅格结构除了可使大量的空间分析模型得以容易实现之外，还具有以下两个特点：（1）易于与遥感相结合。遥感影像是以象元为单位的栅格结构，可以直接将原始数据或经过处理的影像数据纳入栅格结构的地理信息系统。（2）易于信息共享。目前还没有一种公认的矢量结构地图数据记录格式，而不经压缩编码的栅格格式即整数型数据库阵列则易于为大多数程序设计人员和用户理解和使用，因此以栅格数据为基础进行信息共享的数据交流较为实用。

许多实践证明，栅格结构和矢量结构在表示空间数据上可以是同样有效的，对于一个 GIS 软件，较为理想的方案是采用两种数据结构，即栅格结构与矢量结构并存，对于提高地理信息系统的空间分辨率、数据压缩率和增强系统分析、输入输出的灵活性十分重要。两种格式的比较见表 7-2。

表 7-2：矢量格式与栅格格式的比较

	优点	缺点
矢量数据	1．数据结构紧凑、冗余度低 2．有利于网络和检索分析 3．图形显示质量好、精度高	1．数据结构复杂 2．多边形叠加分析比较困难
栅格数据	1．数据结构简单 2．便于空间分析和地表模拟 3．现势性较强	1．数据量大 2．投影转换比较复杂

4．2 相互转换算法

矢量结构与网格结构的相互转换，是地理信息系统的基本功能之一，目前已经发展了许多高效的转换算法；但是，从栅格数据到矢量数据的转换，特别是扫描图像的自动识别，仍然是目前研究的重点。

对于点状实体，每个实体仅由一个坐标对表示，其矢量结构和栅格结构的相互转换基本上只是坐标精度变换问题，不存在太大的技术问题。线实体的矢量结构由一系列坐标对表示，在变为栅格结构时，除把序列中坐标对变为栅格行列坐标外，还需根据栅格精度要求，在坐标点之间插满一系列栅格点，这也容易由两点式直线方程得到。线实体由栅格结构变为矢量结构与将多边形边界表示为矢量结构相似，因此以下重点讨论多边形（面实体）的矢量结构与栅格结构相互转换。

4．2．1 矢量格式向栅格格式的转换

矢量格式向栅格格式转换又称为多边形填充，就是在矢量表示的多边形边界内部的所有栅格点上赋以相应的多边形编码，从而形成类似图 7-4 的栅格数据阵列。几种主要的算法描述如下：

1）内部点扩散算法

该算法由每个多边形一个内部点（种子点）开始，向其八个方向的邻点扩散，判断各个新加入点是否在多边形边界上，如果是边界上，则该新加入点不作为种子点，否则把非边界点的邻点作为新的种子点与原有种子点一起进行新的扩散运算，并将该种子点赋以该多边形的编号。重复上述过程直到所有种子点填满该多边形并遇到边界停止为止。扩散算法程序设计比较复杂，并且在一定的栅格精度上，如果复杂图形的同一多边形的两条边界落在同一个或相邻的两个栅格内，会造成多边形不连通，这样一个种子点不能完成整个多边形的填充。

2）复数积分算法

对全部栅格阵列逐个栅格单元地判断该栅格归属的多边形编码，判别方法是由待判点对每个多边形的封闭边界计算复数积分，对某个多边形，如果积分值为 2πr，则该待判点属于此多边形，赋以多边形编号，否则在此多边形外部，不属于该多边形。

3）射线算法和扫描算法

射线算法可逐点判断数据栅格点在某多边形之外或在多边形内，由待判点向图外某点引射线，判断该射线与某多边形所有边界相交的总次数，如相交偶数次，则待判点在该多边形外部，如为奇数次，则待判点在该多边形内部（图 7-12）。采用射线算法，要注意的是：射线与多边形边界相交时，有一些特殊情况会影响交点的个数，必须予以排除（图 7-13）。

扫描算法是射线算法的改进，将射线改为沿栅格阵列列或行方向扫描线，判断与射线算法相似。扫描算法省去了计算射线与多边形边界交点的大量运算，大大提高了效率。

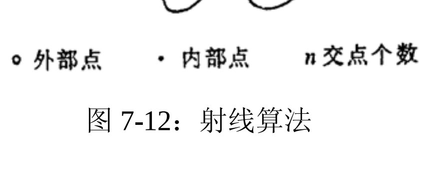


图 7-12：射线算法

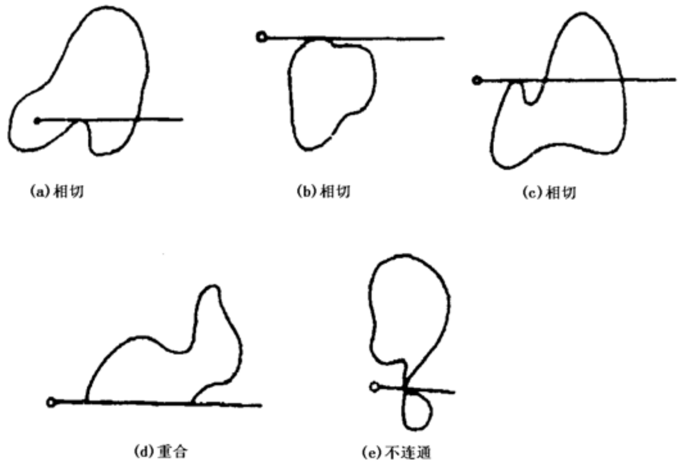


图 7-13：射线算法的特殊情况

4）边界代数算法（BAF-Boundary Algebra Filling）[任伏虎]

边界代数多边形填充算法是一种基于积分思想的矢量格式向栅格格式转换算法，它适合于记录拓扑关系的多边形矢量数据转换为栅格结构。图 7-15 表示转换单个多边形的情况，多边形编号为 a，模仿积分求多边形区域面积的过程，初始化的栅格阵列各栅格值为零，以栅格行列为参考坐标轴，由多边形边界上某点开始顺时针搜索边界线，当边界上行时（图 7-15-a），位于该边界左侧的具有相同行坐标的所有栅格被减去 a；当边界下行时（图 7-15-b），该边界左边（前进方向看为右侧）所有栅格点加一个值 a，边界搜索完毕则完成了多边形的转换。

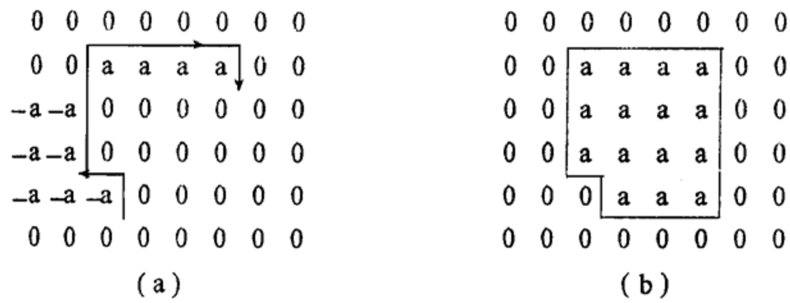


图 7-15：单个多边形的转换

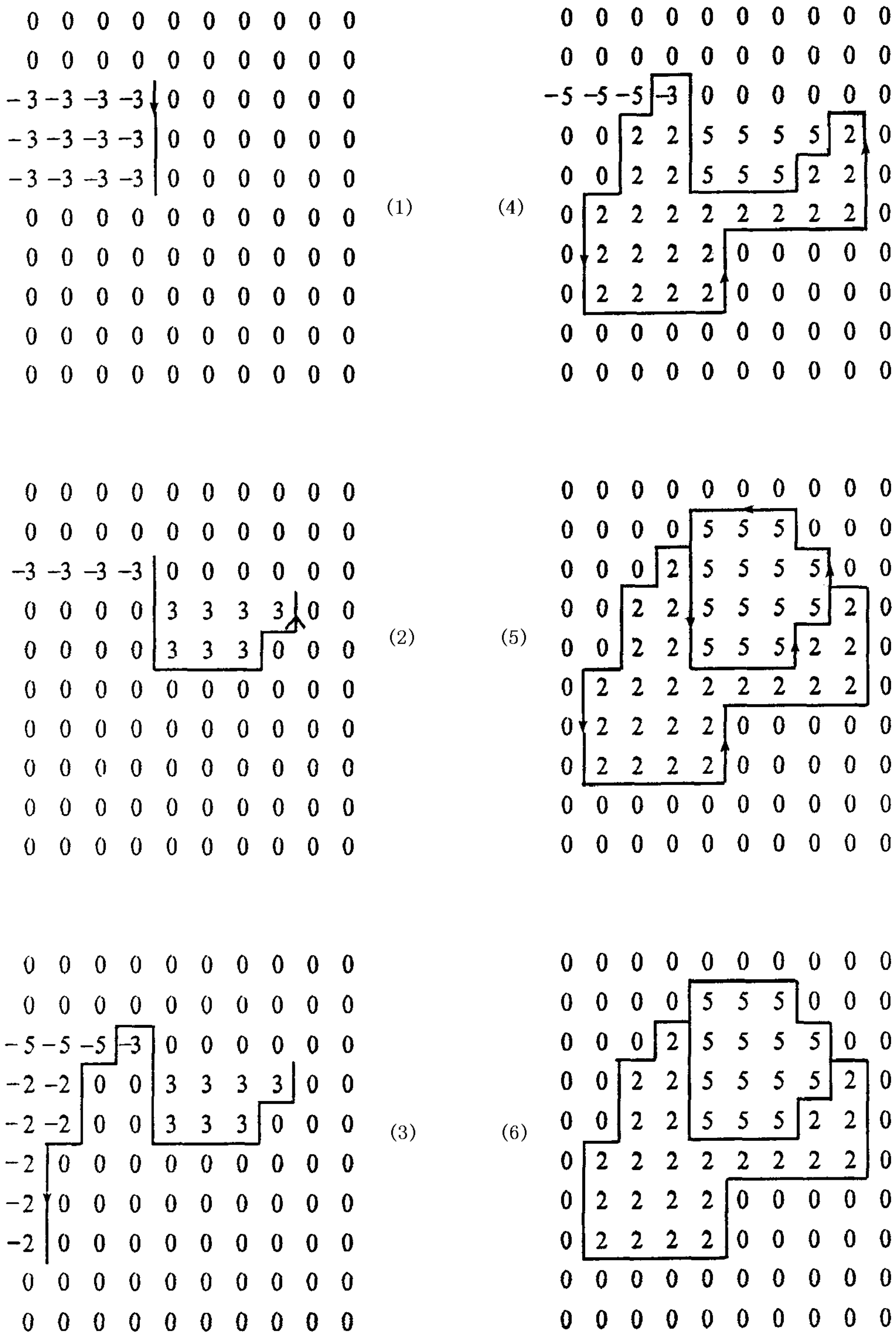


图 7-16：多个多边形的转换

事实上，每幅数字地图都是由多个多边形区域组成的，如果把不属于任何多边形的区域（包含无穷远点的区域）看成编号为零的特殊的多边形区域，则图上每一条边界弧段都与两个不同编号的多边形相邻，按弧段的前进方向分别称为左、右多边形，可以证明，对于这种多个多边形的矢量向栅格转换问题，只需对所有多边形边界弧段作如下运算而不考虑排列次序：当边界弧段上行时，该弧段与左图框之间栅格增加一个值（左多边形编号减去右多边形编号）；当边界弧段下行时，该弧段与左图框之间栅格增加一个值（右多边形编号减去左多边形编号）。两个多边形转换过程如图 7-16 所示。

边界代数法与前述其他算法的不同之处，在于它不是逐点判断与边界的关系完成转换，而是根据边界的拓扑信息，通过简单的加减代数运算将边界位置信息动态地赋给各栅格点，

实现了矢量格式到栅格格式的高速转换，而不需要考虑边界与搜索轨迹之间的关系，因此算法简单、可靠性好，各边界弧段只被搜索一次，避免了重复计算。

但是这并不意味着边界代数法可以完全替代其它算法，在某些场合下，还是要采用种子填充算法和射线算法，前者应用于在栅格图像上提取特定的区域；后者则可以进行点和多边形关系的判断。

4．2．2 栅格格式向矢量格式的转换

多边形栅格格式向矢量格式转换就是提取以相同的编号的栅格集合表示的多边形区域的边界和边界的拓扑关系，并表示由多个小直线段组成的矢量格式边界线的过程。

1）步骤

栅格格式向矢量格式转换通常包括以下四个基本步骤：

（1．1）多边形边界提取：采用高通滤波将栅格图像二值化或以特殊值标识边界点；

（1．2）边界线追踪：对每个边界弧段由一个结点向另一个结点搜索，通常对每个已知边界点需沿除了进入方向的其他 7 个方向搜索下一个边界点，直到连成边界弧段；

（1．3）拓扑关系生成：对于矢量表示的边界弧段数据，判断其与原图上各多边形的空间关系，以形成完整的拓扑结构并建立与属性数据的联系；

（1．4）去除多余点及曲线圆滑：由于搜索是逐个栅格进行的，必须去除由此造成的多余点记录，以减少数据冗余；搜索结果，曲线由于栅格精度的限制可能不够圆滑，需采用一定的插补算法进行光滑处理，常用的算法有：线形迭代法；分段三次多项式插值法；正轴抛物线平均加权法；斜轴抛物线平均加权法；样条函数插值法。

2）多边形栅格转矢量的双边界搜索算法（DBDF-Double Boundary Direct Finding）

算法的基本思想是通过边界提取，将左右多边形信息保存在边界点上，每条边界弧段由两个并行的边界链组成，分别记录该边界弧段的左右多边形编号。边界线搜索采用 2*2 栅格窗口，在每个窗口内的四个栅格数据的模式，可以唯一地确定下一个窗口的搜索方向和该弧段的拓扑关系，极大地加快了搜索速度，拓扑关系也很容易建立。具体步骤如下：

（2．1）边界点和结点提取：采用 2*2 栅格阵列作为窗口顺序沿行、列方向对栅格图像全图扫描，如果窗口内四个栅格有且仅有两个不同的编号，则该四个栅格表示为边界点；如果窗口内四个栅格有三个以上不同编号，则标识为结点（即不同边界弧段的交汇点），保持各栅格原多边形编号信息。对于对角线上栅格两两相同的情况，由于造成了多边形的不连通，也当作结点处理。图 7-17 和图 7-18 给出了节点和边界点的各种情形。

a	b
c	d

(a)

a	a
b	c

(b)

a	b
c	c

(c)

a	b
a	c

(d)

a	b
c	b

(e)

a	b
c	a

(f)

a	b
b	c

(e)

a	b
b	a

(g)

图 7-17：节点的 8 种情形

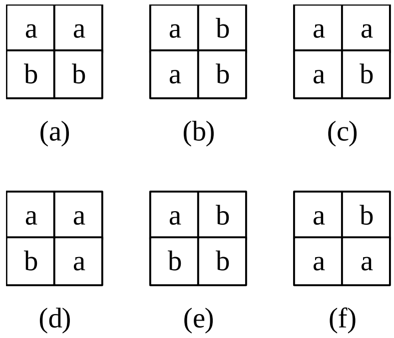


图 7-18：边界点的 6 种情形

（2．2）边界线搜索与左右多边形信息记录：边界线搜索是逐个弧段进行的，对每个弧段由一组已标识的四个结点开始，选定与之相邻的任意一组四个边界点和结点都必定属于某一窗口的四个标识点之一。首先记录开始边界点的两个多边形编号，作为该弧段的左右多边形，下一点组的搜索方向则由进入当前点的搜索方向和该点组的可能走向决定，每个边界点组只能有两个走向，一个是前点组进入的方向，另一个则可确定为将要搜索后续点组的方向。例如图 7-18（c）所示边界点组只可能有两个方向，即下方和右方，如果该边界点组由其下方的一点组被搜索到，则其后续点组一定在其右方；反之，如果该点在其右方的点组之后被搜索到（即该弧段的左右多边形编号分别为 b 和 a），对其后续点组的搜索应确定为下方，其他情况依此类推。可见双边界结构可以唯一地确定搜索方向，从而大大地减少搜索时间，同时形成的矢量结构带有左右多边形编号信息，容易建立拓扑结构和与属性数据的联系，提高转换的效率。

（2．3）多余点去除：多余点的去除基于如下思想：在一个边界弧段上的连续的三个点，如果在一定程度上可以认为在一条直线上（满足直线方程），则三个点中间一点可以被认为上多余的，予以去除。多余点是由于栅格向矢量转换时逐点搜索边界造成的（当边界为直线时），多余点去除算法可大量去除多余点，减少数据冗余。

5．空间索引机制

5．1 索引概念

栅矢一体化空间数据结构一个重要的研究领域是如何建立有效的空间索引结构。目前对线要素索引结构研究较多，主要有 PMR 四叉树、带树和桶方法等，而面要素的索引结构主要有四叉树和 R 树等。这些结构各有自己的应用领域和相对优势，同时也都存在着不足。

空间索引就是指依据空间对象的位置和形状或空间对象之间的某种空间关系按一定的顺序排列的一种数据结构，其中包含空间对象的概要信息，如对象的标识、外接矩形及指向空间对象实体的指针。作为一种辅助性的空间数据结构，空间索引介于空间操作算法和空间对象之间，它通过筛选作用，大量与特定空间操作无关的空间对象被排除，从而提高空间操作的速度和效率。空间索引的性能的优劣直接影响空间数据库和地理信息系统的整体性能，它是空间数据库和地理信息系统的一项关键技术。常见大空间索引一般是自顶向下、逐级划分空间的各种数据结构空间索引，比较有代表性的包括 BSP 树、K-D-B 树、R 树、R+树和 CELL 树等。此外，结构较为简单的格网型空间索引有着广泛的应用。

5．2 索引类型

5．2．1 格网型空间索引

格网型空间索引思路比较简单了，容易理解和实现。其基本思想是将研究区域用横竖线条划分大小相等和不等的格网，记录每一个格网所包含的空间实体。当用户进行空间查询时，首先计算出用户查询对象所在格网，然后再在该网格中快速查询所选空间实体，这样一来就大大地加速了空间索引的查询速度。

5．2．2 BSP 树空间索引

BSP 树是一种二叉树，它将空间逐级进行一分为二的划分（图 7-19）。BSP 树能很好地与空间数据库中空间对象的分布情况相适应，但对一般情况而言，BSP 树深度较大，对各种操作均有不利影响。

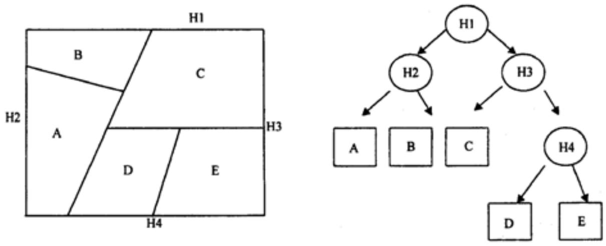


图 7-19：BSP 树

5．2．3 KDB 树空间索引

KDB 树是 B 树向多维空间的一种发展。它对于多维空间中的点进行索引具有较好的动态特性，删除和增加空间点对象也可以很方便地实现；其缺点是不直接支持占据一定空间范围的地物要素，如二维空间中的线和面。该缺点可以通过空间映射或变换的方法部分地得到解决。空间映射或变换就是将 2n 维空间中的区域变换到 2n 维空间中的点，这样便可利用点索引结构来对区域进行索引，原始空间的区域查询便转化为高维空间的点查询。但空间映射或变换方法仍然存在着缺点：高维空间的点查询要比原始空间的点查询困难得多；经过变换，原始空间中相邻的区域有可能在点空间中距离变得相当遥远，这些都将影响空间索引的性能。

5．2．4 R 树和 R+树

R 树根据地物的最小外包矩形建立（图 7-20），可以直接对空间中占据一定范围的空间对象进行索引。R 树的每一个结点 N 都对应着磁盘页 D（N）和区域 I（N），如果结点不是叶结点，则该结点的所有子结点的区域都在区域 I（N）的范围之内，而且存储在磁盘页 D（N）中；如果结点是叶结点，那么磁盘页 D（N）中存储的将是区域 I（N）范围内的一系列子区域，子区域紧紧围绕空间对象，一般为空间对象的外接矩形。

R 树中每个结点所能拥有的子结点数目是有上下限的。下限保证索引对磁盘空间的有效利用，子结点的数目小于下限的结点将被删除，该结点的子结点将被分配到其它的结点中；设立上限的原因是因为每一个结点只对应一个磁盘页，如果某个结点要求的空间大于一个磁盘页，那么该结点就要被划分为两个新的结点，原来结点的所有子结点将被分配到这两个新的结点中。

由于 R 树兄弟结点对应的空间区域可以重叠，因此，R 树可以较容易地进行插入和删除操作；但正因为区域之间有重叠，空间索引可能要对多条路径进行搜索后才能得到最后的结果，因此，其空间搜索的效率较低。正是这个原因促使了 R+树（图 7-21）的产生。在 R+树中，兄弟结点对应的空间区域没有重叠，而没有重叠的区域划分可以使空间索引搜索的速度大大提高；但由于在插入和删除空间对象时要保证兄弟结点对应的空间区域不重叠，而使插入和删除操作的效率降低。

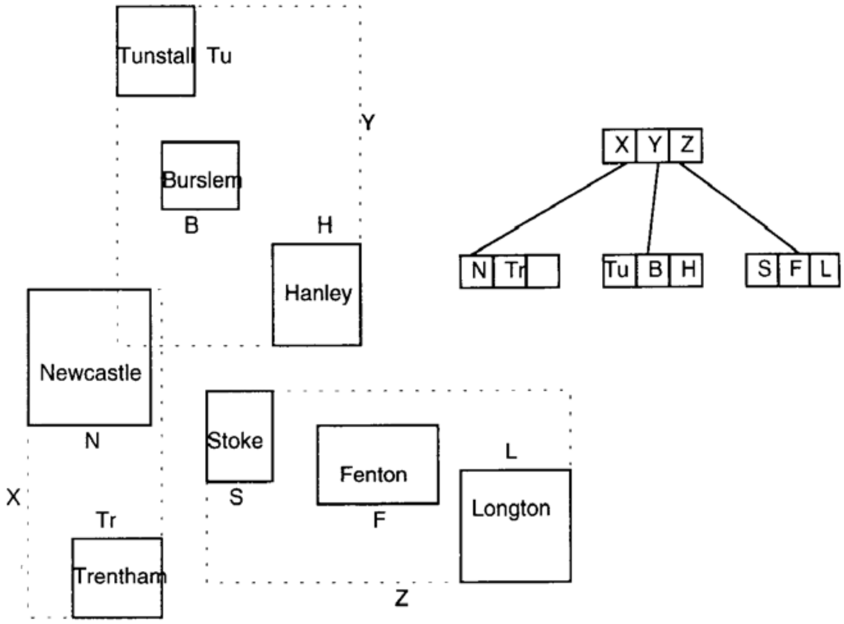


图 7-20：R 树

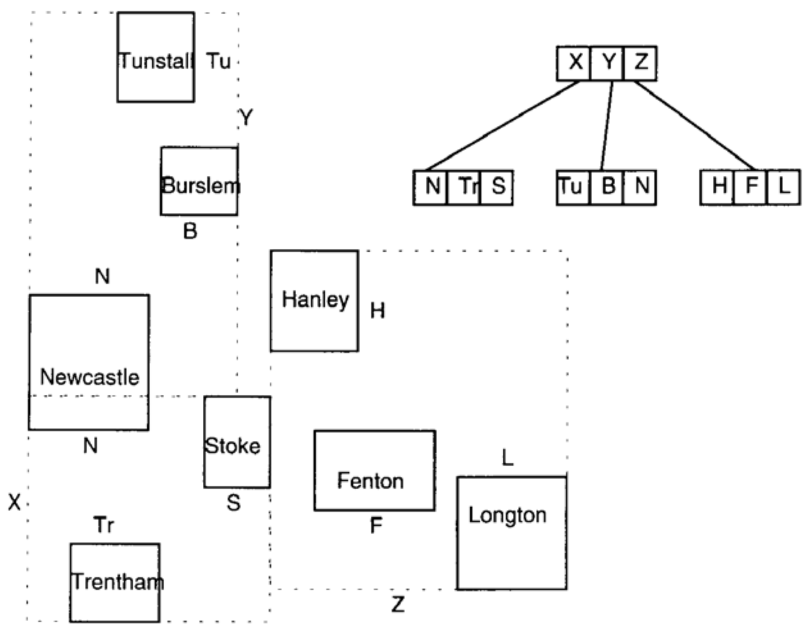


图 7-21：R+树

5．2．5 CELL 树

考虑到 R 树和 R+在插入、删除和空间搜索效率两方面难于兼顾，CELL 树应运而生。它在空间划分时不再采用矩形作为划分的基本单位，而是采用凸多边形来作为划分的基本单位，具体划分方法与 BSP 树有类似之处，子空间不再相互覆盖。CELL 树的磁盘访问次数比 R 树和 R+树少，由于磁盘访问次数是影响空间索引性能的关键指标，故 CELL 树是比较优秀的空间索引方法（图 7-22）。

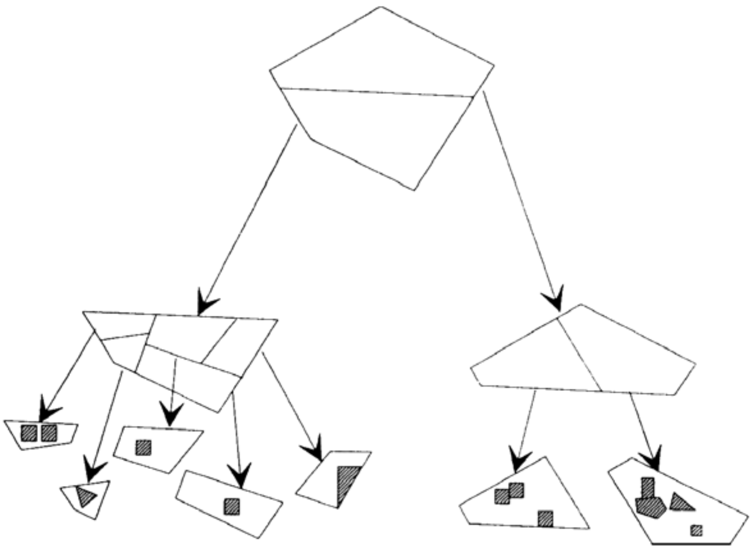


图 7-22：CELL 树

6．空间信息查询

目前大多数成熟的商品化地理信息系统软件的查询功能都可完美地实现对空间实体的简单查找，如根据鼠标所指的空间位置，系统可查找出该位置的空间实体和空间范围（由若干个空间实体组成）以及它们的属性，并显示出该空间对象的属性列表，并可以进行有关统计分析。该查询工作可分为两步：首先借助于空间索引，在空间数据库中快速检索出被选空间实体；然后，根据空间数据和属性数据的连接即可得到该空间实体的属性列表。

6．1 基于属性特征查询

一般来说，基于属性信息的查询操作主要是在属性数据库中完成的。目前大多数的地理信息系统软件都将属性信息存储在关系数据库中，而发展成熟的关系数据库又为我们提供了完备的数据索引方法及信息查询手段。几乎所有的关系数据库管理系统都支持标准的结构化查询语言。利用 SQL，我们可以在属性数据库中很方便地实现属性信息的复合条件查询，筛选出满足条件的空间实体的标识值，再到空间数据库中根据标识值检索到该空间实体。

6．2 基于空间关系和属性特征的查询（SQL）

空间实体间有着许多空间关系（包括拓扑、顺序、度量等关系）。在实际应用过程中，用户往往希望地理信息系统提供一些更能直接计算空间实体关系的功能，如用户希望查询出满足如下条件的城市：

A 在某条铁路的东部；B 距离该铁路不超过 30 公里；C 城市人口大于 70 万；D 城市选择区域是特定的多边形。整个查询计算涉及了空间顺序关系（铁路东部）、空间距离关系（距离该铁路不超过 30 公里）、空间拓扑关系（被选城市在特定的选择区域之内）、属性信息查询（城市人口大于 70 万）。

就目前成熟的地理信息系统而言，比较系统地完成上述查询任务还较为困难。为此，众多的地理信息系统专家提出了“空间查询语言”（Spatial Query Language）以作为解决问题的方案，但仍处于理论发展和技术探索阶段。

6．3 一种空间扩展 SQL 查询语言——GeoSQL

查询、检索是地理信息系统中使用最频繁的功能之一，GIS 用户提出的大部分问题都可以表达为查询形式，即空间查询语言不仅可以使 GIS 用户方便地访问、查询和处理空间数据，也可以实现空间数据的安全性和完整性控制。相对于一般 SQL，空间扩展 SQL 主要增加了空间数据类型和空间操作算子，以满足空间特征的查询。空间特征包含空间属性和非空间属性，空间属性由特定的“Location”字段来表示。空间数据类型除具有一般的整型、实型、字符串外，还具有下列空间数据类型：点类型、弧段类型、不封闭的线类型、（Polygon）多边形类型、图像类型、复杂空间特征类型。以上类型是针对“位置”字段而言的。GeoSQL 中的空间操作算子是指带有参数的函数。通常它以空间特征为参数，返回空间特征或数值。空间操作算子主要分为两类：一元空间操作算子和二元空间操作算子。通常，标准 SQL 的一般形式为：SELECT…FROM…WHERE，分别对应关系操作投影、笛卡尔积和选择，其中，FROM 语句代表所给关系的笛卡尔积，也就是定义了一个单独的关系。WHERE 语句中的选择和 SELECT 语句中的投影均作用于该关系上。尽管 GeoSQL 属于非过程化文本语言，但为使操作简单方便，它仍可以借鉴可视化查询语言的特点，即使用图符、列表框等组件来尽量减少用户的文本输入，同时也可防止输入时由于误记、语义误解等产生的语法错误。GeoSQL 的实现过程如图 7-23 所示。

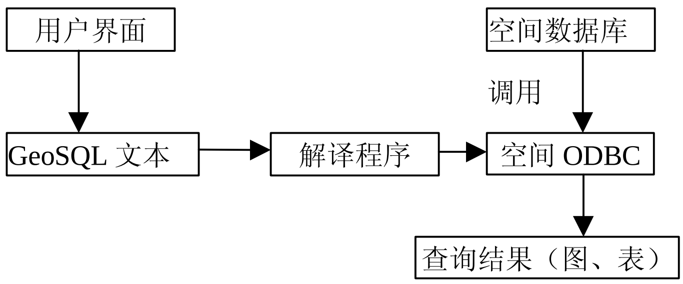


图 7-23：GeoSQL 的实现过程

地理学中的第一条法则，任何事物都与其它事情相关，但是距离近的事物比距离远的关系更大。

Waldo Tobler

第八章 空间分析

导读：空间分析源于 60 年代地理和区域科学的计量革命，在开始阶段，主要是应用定量（主要是统计）分析手段用于分析点、线、面的空间分布模式。后来更多的是强调地理空间本身的特征、空间决策过程和复杂空间系统的时空演化过程分析。实际上自有地图以来，人们就始终在自觉或不自觉地进行着各种类型的空间分析。如在地图上量测地理要素之间的距离、方位、面积，乃至利用地图进行战术研究和战略决策等，都是人们利用地图进行空间分析的实例，而后者实质上已属较高层次上的空间分析。

地理信息系统集成了多学科的最新技术，如关系数据库管理，高效图形算法，插值，区划和网络分析，为空间分析提供了强大的工具，使得过去复杂困难的高级空间分析任务变得简单易行。目前绝大多数地理信息系统软件都有空间分析功能。空间分析早已成为地理信息系统的核心功能之一，它特有的对地理信息（特别是隐含信息）的提取、表现和传输功能，是地理信息系统区别于一般信息系统的主要功能特征。

空间分析是对分析空间数据有关技术的统称。根据作用的数据性质不同，可以分为：（1）基于空间图形数据的分析运算；（2）基于非空间属性的数据运算；（3）空间和非空间数据的联合运算。空间分析赖以进行的基础是地理空间数据库，其运用的手段包括各种几何的逻辑运算、数理统计分析，代数运算等数学手段，最终的目的是解决人们所涉及到地理空间的实际问题，提取和传输地理空间信息，特别是隐含信息，以辅助决策。

本章介绍 GIS 中实现空间分析的基本功能，包括空间查询与量算，缓冲区分析、叠加分析、路径分析、空间插值、统计分类分析等，并描述了相关的算法，以及其中的计算公式。

1. 空间查询与量算

查询和定位空间对象，并对空间对象进行量算是地理信息系统的基本功能之一，它是地理信息系统进行高层次分析的基础。在地理信息系统中，为进行高层次分析，往往需要查询定位空间对象，并用一些简单的量测值对地理分布或现象进行描述，如长度，面积，距离，形状等。实际上，空间分析首先始于空间查询和量算，它是空间分析的定量基础。

1．1 空间查询

图形与属性互查是最常用的查询，主要有两类：第一类是按属性信息的要求来查询定位空间位置，称为“属性查图形”。如在中国行政区划图上查询人口大于 4000 万且城市人口大于 1000 万的省有哪些，这和一般非空间的关系数据库的 SQL 查询没有区别，查询到结果后，再利用图形和属性的对应关系，进一步在图上用指定的显示方式将结果定位绘出。第二类是根据对象的空间位置查询有关属性信息，称为“图形查属性”。如一般地理信息系统软件都

提供一个“INFO”工具，让用户利用光标，用点选、画线、矩形、圆、不规则多边形等工具选中地物，并显示出所查询对象的属性列表，可进行有关统计分析。该查询通常分为两步，首先借助空间索引，在地理信息系统数据库中快速检索出被选空间实体，然后根据空间实体与属性的连接关系即可得到所查询空间实体的属性列表。

在大多数 GIS 中，提供的空间查询方式有：

1）基于空间关系查询

空间实体间存在着多种空间关系，包括拓扑、顺序、距离、方位等关系。通过空间关系查询和定位空间实体是地理信息系统不同于一般数据库系统的功能之一。如查询满足下列条件的城市：

- *在京沪线的东部*
- *距离京沪线不超过 50 公里*
- *城市人口大于 100 万*
- *城市选择区域是特定的多边形；*

整个查询计算涉及了空间顺序方位关系（京沪线东部），空间距离关系（距离京沪线不超过 50 公里），空间拓扑关系（使选择区域是特定的多边形），甚至还有属性信息查询（城市人口大于 100 万）。

简单的面、线、点相互关系的查询包括：

- 面面查询，如与某个多边形相邻的多边形有哪些。
- 面线查询，如某个多边形的边界有哪些线。
- 面点查询，如某个多边形内有哪些点状地物。
- 线面查询，如某条线经过（穿过）的多边形有哪些，某条链的左、右多边形是哪些。
- 线线查询，如与某条河流相连的支流有哪些，某条道路跨过哪些河流。
- 线点查询，如某条道路上有哪些桥梁，某条输电线上有哪些变电站。
- 点面查询，如某个点落在哪个多边形内。
- 点线查询，如某个结点由哪些线相交而成。

2）基于空间关系和属性特征查询

传统的关系数据库的标准 SQL 并不能处理空间查询，这是由于关系数据库技术的弱点造成的，对于 GIS 而言，需要对 SQL 进行扩展。对于传统的 SQL，要实现空间操作，需要将 SQL 命令嵌入一种编程语言中，如 C 语言；而新的 SQL 允许用户定义自己的操作，并嵌入到 SQL 命令中。

3）地址匹配查询

根据街道的地址来查询事物的空间位置和属性信息是地理信息系统特有的一种查询功能，这种查询利用地理编码，输入街道的门牌号码，就可知道大致的位置和所在的街区。它对空间分布的社会、经济调查和统计很有帮助，只要在调查表中添了地址，地理信息系统可以自动地从空间位置的角度来统计分析各种经济社会调查资料。另外这种查询也经常用于公用事业管理，事故分析等方面，如邮政、通讯、供水、供电、治安、消防、医疗等领域。

1．2 空间量算

1．2．1 几何量算

几何量算对不同的点、线、面地物有不同的含义：

- 点状地物（0 维）：坐标；
- 线状地物（1 维）：长度，曲率，方向；
- 面状地物（2 维）：面积，周长，形状，曲率等；
- 体状地物（3 维）：体积，表面积等。

一般的 GIS 软件都具有对点、线、面状地物的几何量算功能，或者是针对矢量数据结构，或者是针对栅格数据结构的空间数据。

1）线的长度计算

线状地物对象最基本的形态参数之一是长度。在矢量数据结构下，线表示为点对坐标（X，Y）或（X，Y，Z）的序列，在不考虑比例尺情况下，线长度的计算公式为：

$$L=\sum_{i=0}^{n-1}\left[\left(X_{i+1}-X_i\right)^2+\left(Y_{i+1}-Y_i\right)^2+\left(Z_{i+1}-Z_i\right)^2\right]^{1/2}=\sum_{i=1}^{n}l_i$$

对于复合线状地物对象，则需要在对诸分支曲线求长度后，再求其长度总和。

通过离散坐标点对串来表达线对象，选择反映曲线形状的选点方案非常重要，往往由于选点方案不同，会带来长度计算的不同精度问题。为提高计算精度，增加点的数目，会对数据获取、管理与分析带来额外的负担，折中的选点方案是在曲线的拐弯处加大点的数目，在平直段减少点数，以达到计算允许精度要求。

在栅格数据结构里，线状地物的长度就是累加地物骨架线通过的格网数目，骨架线通常采用 8 方向连接，当连接方向为对角线方向时，还要乘上$\sqrt{2}$。

2）面状地物的面积

面积是面状地物最基本的参数。在矢量结构下，面状地物以其轮廓边界弧段构成的多边形表示的。对于没有空洞的简单多边形，假设有 N 个顶点，其面积计算公式为：

$$S=\left|\frac{1}{2}\left(\sum_{i=1}^{N-2}\left(x_i y_{i+1}-x_{i+1}y_i\right)+\left(x_N y_1-x_1 y_N\right)\right)\right|$$

所采用的是几何交叉处理方法，即沿多边形的每个顶点作垂直与 X 轴的垂线，然后计算每条边、它的两条垂线及这两条垂线所截得 X 轴部分所包围的面积，所求出的面积的代数和，即为多边形面积。对于有孔或内岛的多边形，可分别计算外多边形与内岛面积，其差值为原多边形面积。此方法亦适合于体积的计算。

对于栅格结构，多边形面积计算就是统计具有相同属性值的格网数目。但对计算破碎多边形的面积有些特殊，可能需要计算某一个特定多边形的面积，必须进行再分类，将每个多边形进行分割赋给单独的属性值，之后再进行统计。

1．2．2 形状量算

面状地物形状量测的两个基本考虑：空间一致性问题，即有孔多边形和破碎多边形的处理；多边形边界特征描述问题。

度量空间一致性最常用的指标是欧拉函数，用来计算多边形的破碎程度和孔的数目。欧拉函数的结果是一个数，称为欧拉数。欧拉函数的计算公式为：

欧拉数=（孔数）-（碎片数-1）

图 8-1 表示了多边形的三种可能的情形。

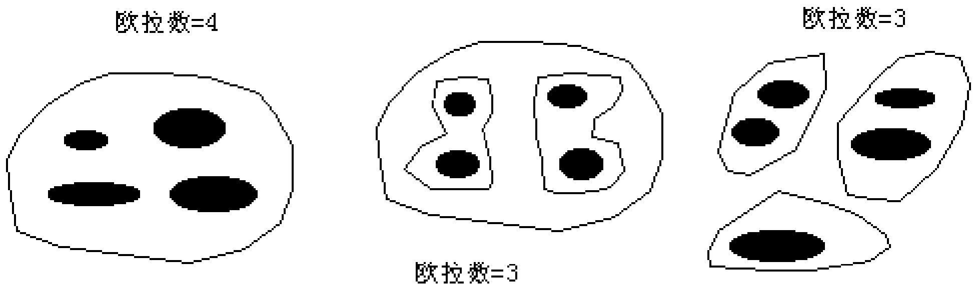


图 8-1：欧拉数

对于图(a)，欧拉数=4-(1-1)=4 或欧拉数=4-0=4；对于图(b)欧拉数=4-(2-1)=3 或欧拉数=4-1=3；图(c)欧拉数=5-(3-1)=3。

关于多边形边界描述的问题，由于面状地物的外观是复杂多变的，很难找到一个准确的指标进行描述。最常用的指标包括多边形长、短轴之比，周长面积比，面积长度比等。其中绝大多数指标是基于面积和周长的。通常认为圆形地物既非紧凑型也非膨胀型，则可定义其形状系数 r 为：

$$r = \frac{P}{2\sqrt{\pi} \cdot \sqrt{A}}$$

其中 P 为地物周长，A 为面积。如果 r<1 为紧凑型；r=1 为标准圆；r>1 为膨胀型。

1．2．3 质心量算

质心是描述地理对象空间分布的一个重要指标。例如要得到一个全国的人口分布等值线图，而人口数据只能到县级，所以必须在每个县域里定义一个点作为质心，代表该县的数值，然后进行插值计算全国人口等值线。质心通常定义为一个多边形或面的几何中心，当多边形比较简单，比如矩形，计算很容易。但当多边形形状复杂时，计算也更加复杂。

在某些情况下，质心描述的是分布中心，而不是绝对几何中心。同样以全国人口为例，当某个县绝大部分人口明显集中于一侧时，可以把质心放在分布中心上，这种质心称为平均中心或重心。如果考虑其它一些因素的话，可以赋予权重系数，称为加权平均中心。计算公式是：

$$X_G = \frac{\sum_i W_i X_i}{\sum_i W_i}$$

$$Y_G = \frac{\sum_i W_i Y_i}{\sum_i W_i}$$

其中，W_i为第i个离散目标物权重，X_i，Y_i为第i个离散目标物的坐标。

质心量测经常用于宏观经济分析和市场区位选择，还可以跟踪某些地理分布的变化，如人口变迁，土地类型变化等。

1．2．4 距离量算

“距离”是人们日常生活中经常涉及到的概念，它描述了两个事物或实体之间的远近程度。最常用的距离概念是欧氏距离，无论是矢量结构，还是栅格结构都很容易实现。在 GIS 中，距离通常是两个地点之间的计算，但有时人们想知道一个地点到所有其它地点的距离，这时得到的距离是一个距离表面。如果一区域中所有的性质与方向无关，则称为各向同性区域。以旅行时间为例，如果从某一点出发，到另一点的所耗费的时间只与两点之间的欧氏距离成正比，则从一固定点出发，旅行特定时间后所能达到的点必然组成一个等时圆。而现实生活中，旅行所耗费的时间不只与欧氏距离成正比，还与路况、运输工具性能等有关，从固定点出发，旅行特定时间后所能到达的点则在各个方向上是不同距离的，形成各向异性距离表面。（图 8-2）

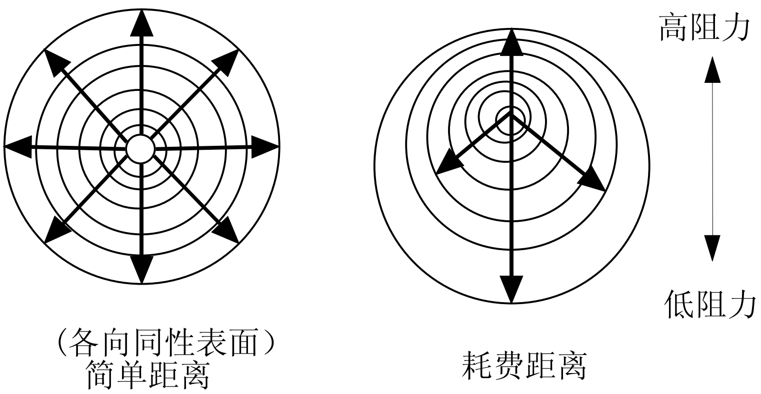


图 8-2：各向同性和各向异性的距离表面

考虑到阻力影响，计算的距离称为耗费距离。物质在空间中移动总要花费一些代价，如资金、时间等。阻力越大耗费也越大。相应的通过耗费距离得到的距离表面称为阻力表面或耗费表面，其属性值代表一耗费或阻力大小。可以根据阻力表面计算最小耗费距离。

对于描述点、线、面坐标的矢量结构，也有一系列的不同于欧氏距离的概念。欧氏距离通常用于计算两点的直线距离：

$$d = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}$$

当有障碍或阻力存在时，两点之间的距离就不能用直线距离，计算非标准欧氏距离的一般公式为：

$$d = \left[(X_i - X_j)^k + (Y_i - Y_j)^k\right]^{1/k}$$

当 *k=2* 时，就是欧氏距离计算公式。当 *k=1* 时，得到的距离称为曼哈顿距离。欧氏距离、曼哈顿距离和非欧氏距离的计算如图 8-3 所示。

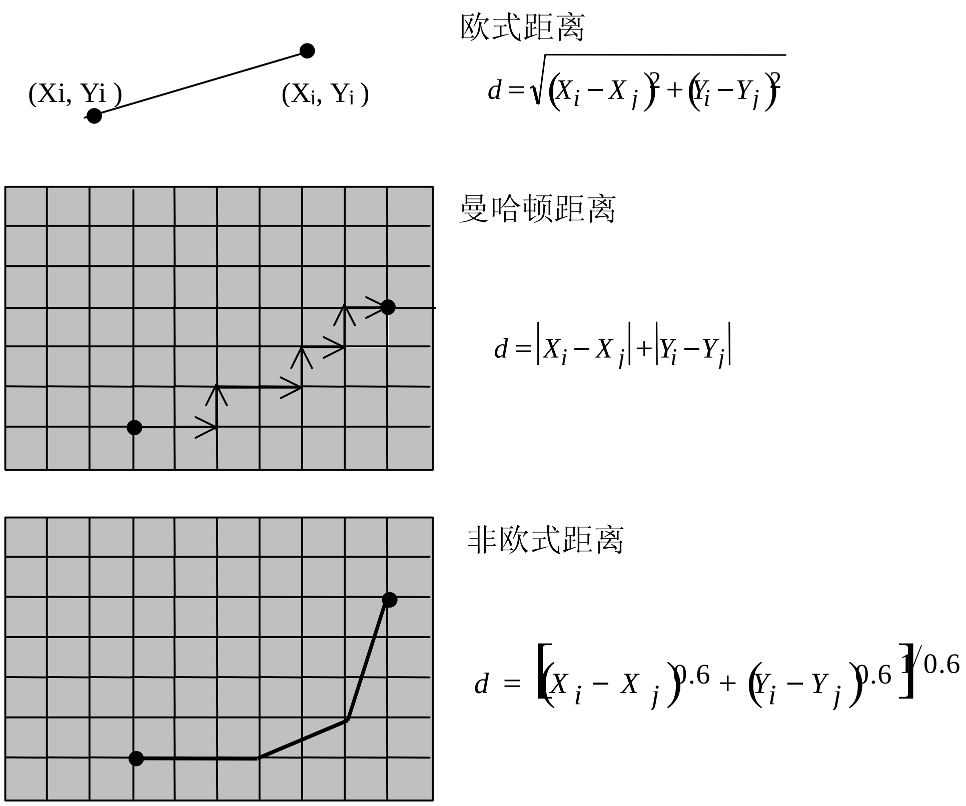


图 8-3：欧氏距离、曼哈顿距离和一种非欧氏距离

2．空间变换

地理信息系统通常是按有一定意义的图层和相应的属性建立空间数据库的。为了满足特定空间分析的需要，需对原始图层及其属性进行一系列的逻辑或代数运算，以产生新的具有特殊意义的地理图层及其属性，这个过程称为空间变换。空间变换可以基于单个图层进行，也可以对多个图层，本章将空间变换仅限于对单个图层的操作或计算，基于多图层的操作，将在叠加分析里讲述。

地理信息系统中空间数据可分为矢量和栅格两种数据结构。由于矢量结构中包含了大量的拓扑信息，数据组织复杂，使得空间变换十分繁琐。而栅格结构简单规则，空间变换比较容易。另外基于矢量结构的空间变换，对于单个图层意义不大，生成新图层时往往需要多个图层的信息，在多图层叠加分析中意义很大。

基于栅格结构的空间变换可分为三种方式：（1）单点变换；（2）邻域变换；（3）区域变换。

单点变换只考虑单个点的属性值进行运算，假定独立单元的变换不依赖于其邻点上属性的影响，也不受区域内一般特征的影响。单点变换最常见的函数有加、减、乘、除等代数运算；与、并、非、异或等逻辑运算；大于、小于等比较运算；指数函数，对数函数、三角函数等。其得到的新图层可与原图层属性意义完全不同。

邻域变换是指在计算新图层图元值时，不仅考虑原始图层上相应图元本身的值，而且还要考虑与该图元有邻域关联的其它图元值的影响。这种关联可以是直接的几何关联，也可能是间接的几何关联。常见的函数有平滑、离散点搜索、连续表面描述（坡度、坡向、可视域分析）、点在多边形中的判断等。

区域变换是指在计算新图层属性值时，要考虑整个区域的属性值，即通过一个函数对某

一区域内的所有值进行综合，然后计算新属性值。常见的函数有求区域平均值、众数，极值、求和、归组、整体插值等方法。

3. 再分类

通过分类找出隐藏信息是地理信息系统的重要功能之一。与传统地图相比，地图上所载负的数据是经过专门分类和处理过的，而地理信息系统存储的数据则具有原始数据的性质，所以可以根据不同的需要对数据再进行分类和提取。由于这种分类是对原始数据进行的再次分类组织，因此称为再分类（Reclassification）。

地理信息系统区别于其它信息系统的方面是其对空间信息的处理功能，同时也提供了对非空间属性的处理功能，尽管比较简单，但它在实际应用中有着重要的作用。根据地理信息的非空间属性，如材料、价值、使用性质等，进行再分类，这种纯粹基于非空间属性的分类，与其它信息系统对简单结构化的数据进行分类的方法是一样的，可以使用经典的数理统计方法，如主成分分析、层次分析、聚类分析、判别分析等等。这种分类属于普通的分类，它不改变地物已有的属性值，而只是根据地物的属性，将它们划分到相应的类别中。本章主要论述 GIS 中通过地物属性信息，经过分类组织产生新地物特征的再分类。

点、线状地物的再分类，对于矢量数据结构可以通过简单的修改属性表中的数值来实现，对于栅格数据结构也可以通过修改属性值来获得新的点、线地物。面状地物的再分类，对于栅格数据结构则和点、线分类一样，简单的改变属性数值并改变图例表现这一变化。例如有一个栅格图，属性值从 1 到 15 分别代表一种农作物，如果 1 到 5 及 13 为粮食作物，其它代表经济作物，可将 1 到 5 和 13 重新赋值 1，其它数赋值 2，则可得到只有粮食作物和经济作物两类地物的栅格图，并改变图例体现这一变化。对于矢量数据结构的面状地物再分类，则需要同时改变实体的几何形状和属性。首要的任务是去掉将要合并的多边形之间的分界线（Line Dissolve），再把这两个多边形的属性值变为同一属性。如图 8-4 所示：

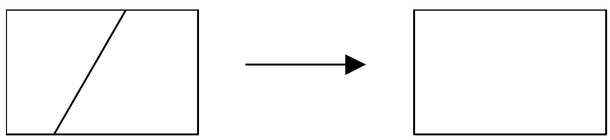


图 8-4：多边形的合并

因为对面状地物的再分类得到的新图层的类别比原图层少，称为归组（Group），它是最常用和最简单的再分类。如果想把面状地物进一步分解成不同类别的地物，就不能用此方法，因为不能知道界线的位置，可使用另一个图层，通过多边形叠加方法来实现*。

上面讲的再分类方法，都是只根据面状地物本身的属性，通过重新改变属性值而实现分类的目的，当然也可以结合邻域范围的属性值进行再分类。如坡度计算，缓冲区计算。再分类还可以综合多个图层的属性信息，如图 8-5 所示。

* 关于多边形叠加将在下面章节论述。

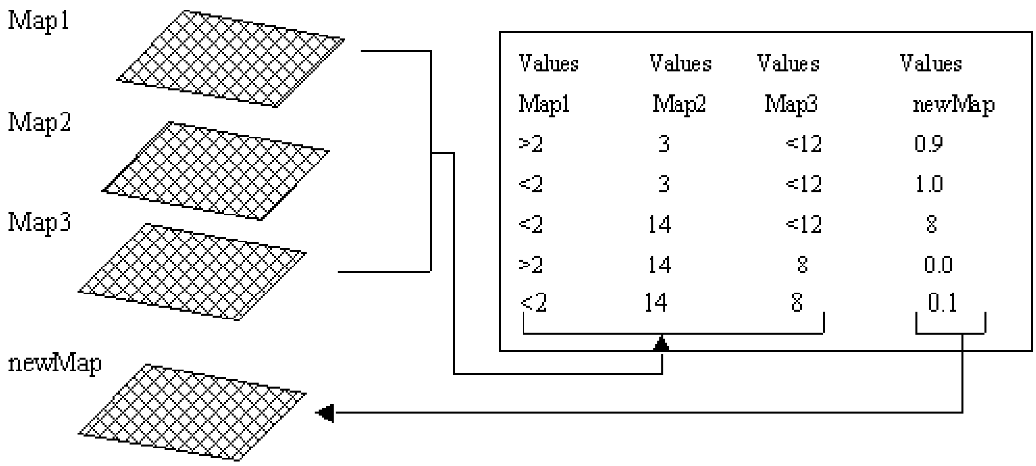


图 8-5：多个属性的再分类

4．缓冲区分析

邻近度（Proximity）描述了地理空间中两个地物距离相近的程度，其确定是空间分析的一个重要手段。交通沿线或河流沿线的地物有其独特的重要性，公共设施（商场，邮局，银行，医院，车站，学校等）的服务半径，大型水库建设引起的搬迁，铁路，公路以及航运河道对其所穿过区域经济发展的重要性等，均是一个邻近度问题。缓冲区分析是解决邻近度问题的空间分析工具之一 。

所谓缓冲区就是地理空间目标的一种影响范围或服务范围。从数学的角度看，缓冲区分析的基本思想是给定一个空间对象或集合，确定它们的邻域，邻域的大小由邻域半径R决定。因此对象O_i的缓冲区定义为：

$$B_i = \{x : d(x, O_i) \le R\}$$

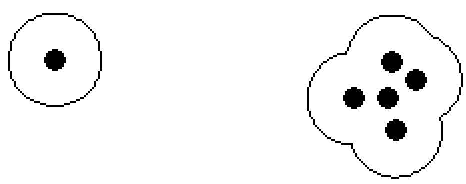
即对象 O_i 的半径为R的缓冲区为距O_i的距离d小于R的全部点的集合。d一般是最小欧氏距离，但也可是其它定义的距离。对于对象集合

$$O = \{O_i : i = 1,2,\cdots,n\}$$

其半径为 R 的缓冲区是各个对象缓冲区的并，即：

$$B = \bigcup_{i=1}^{n} B_i$$

图 8-6 为点对象、线对象、面对象及对象集合的缓冲区示例。



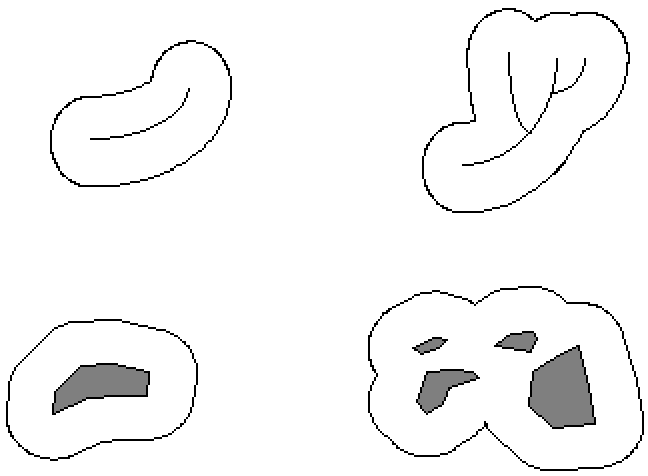


图 8-6：点、线、多边形的缓冲区

另外还有一些特殊形态的缓冲区，如点对象有三角形，矩形和圈形等，对于线对象有双侧对称，双侧不对称或单侧缓冲区，对于面对象有内侧和外侧缓冲区。这些适合不同应用要求的缓冲区，尽管形态特殊，但基本原理是一致的。

缓冲区计算的基本问题是双线问题。双线问题有很多另外的名称，如图形加粗，加宽线，中心线扩张等，它们指的都是相同的操作。

1）角分线法

双线问题最简单的方法是角分线法（简单平行线法）。算法是在轴线首尾点处，作轴线的垂线并按缓冲区半径 R 截出左右边线的起止点；在轴线的其它转折点上，用与该线所关联的前后两邻边距轴线的距离为 R 的两平行线的交点来生成缓冲区对应顶点。如图 8-7 所示。

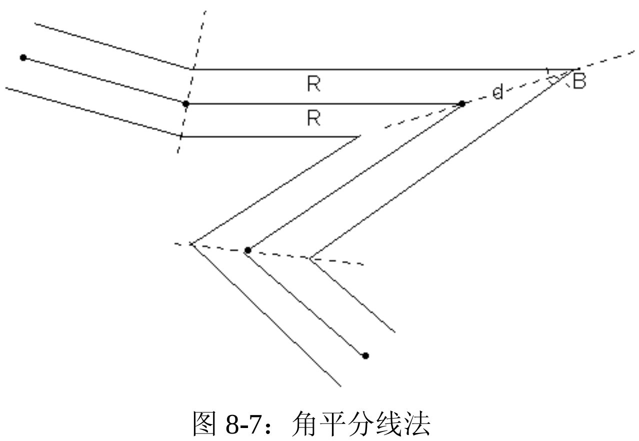


图 8-7：角平分线法

角分线法的缺点是难以最大限度保证双线的等宽性，尤其是在凸侧角点在进一步变锐时，将远离轴线顶点。根据上图，远离情况可由下式表示：

$$d = R/\sin(B/2)$$

当缓冲区半径不变时，d 随张角 B 的减小而增大，结果在尖角处双线之间的宽度遭到破坏。

因此，为克服角分线法的缺点，要有相应的补充判别方案，用于校正所出现的异常情况。但由于异常情况不胜枚举，导致校正措施繁杂。

2）凸角圆弧法

在轴线首尾点处，作轴线的垂线并按双线和缓冲区半径截出左右边线起止点；在轴线其

它转折点处，首先判断该点的凸凹性，在凸侧用圆弧弥合，在凹侧则用前后两邻边平行线的交点生成对应顶点。这样外角以圆弧连接，内角直接连接，线段端点以半圆封闭。如图 8-8 所示。

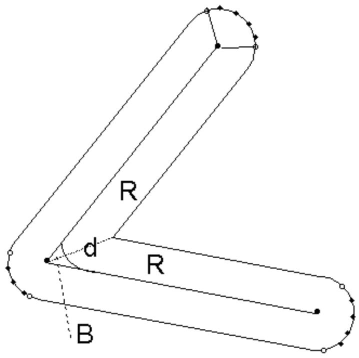


图 8-8：凸角圆弧法

在凹侧平行边线相交在角分线上。交点距对应顶点的距离与角分线法类似公式：

$$d = R/\sin(B/2)$$

该方法最大限度的保证了平行曲线的等宽性，避免了角分线法的众多异常情况。

该算法非常重要的一环是折点凸凹性的自动判断。此问题可转化为两个矢量的叉积：把相邻两个线段看成两个矢量，其方向取坐标点序方向。若前一个矢量以最小角度扫向第二个矢量时呈逆时针方向，则为凸顶点，反之为凹顶点。具体算法过程如下：

由矢量代数可知，矢量 $\overrightarrow{AB}$，$\overrightarrow{BC}$ 可用其端点坐标差表示（9-9）：

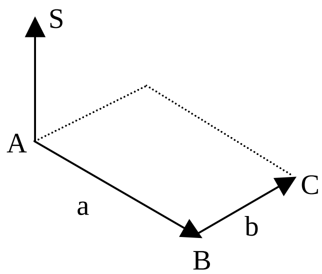


图 8-9：采用向量叉乘判断向量排列

$$\overrightarrow{AB} = (X_B - X_A, Y_B - Y_A) = (a_x, a_y)$$

$$\overrightarrow{BC} = (X_C - X_B, Y_C - Y_B) = (b_x, b_y)$$

$$\vec{S} = \overrightarrow{AB} \times \overrightarrow{BC} = \vec{a} \times \vec{b} = (a_x b_y - b_x a_y)$$
$$= (X_B - X_A)(Y_C - Y_B) - (X_C - X_B)(Y_B - Y_A)$$

矢量代数叉积遵循右手法则，即当 ABC 呈逆时针方向时，S 为正，否则为负。

若 S>0，则 ABC 呈逆时针，顶点为凸；

若 S<0，则 ABC 呈顺时针，顶点为凹；

若 S=0，则 ABC 三点共线。

对于简单情形，缓冲区是一个简单多边形，但当计算形状比较复杂的对象或多个对象集合的缓冲区时，就复杂得多。为使缓冲区算法适应更为普遍的情况，就不得不处理边线自相

交的情况。当轴线的弯曲空间不容许双线的边线无压盖地通过时，就会产生若干个自相交多边形。图 8-10 给出一个缓冲区边线自相交的例子。

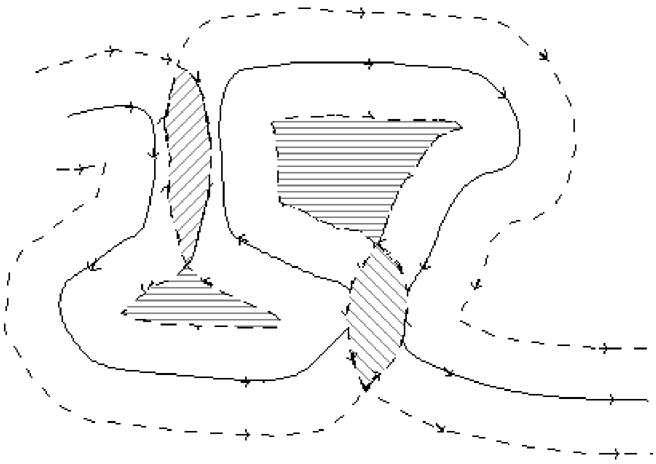


图 8-10：缓冲区边界相交的情况

自相交多边形分为两种情况：岛屿多边形和重叠多边形。岛屿多边形是缓冲区边线的有效组成部分；重叠多边形不是缓冲区边线的有效组成，不参与缓冲区边线的最终重构。对于岛屿多边形和重叠多边形的自动判别方法，首先定义轴线坐标点序为其方向，缓冲区双线分成左右边线，左右边线自相交多边形的判别情形恰好对称。对于左边线，岛屿自相交多边形呈逆时针方向，重叠自相交多边形呈顺时针方向；对于右边线，岛屿多边形呈顺时针方向，重叠多边形呈逆时针方向。

当存在岛屿和重叠自相交多边形时，最终计算的边线被分为外部边线和若干岛屿。对于缓冲区边线绘制，只要把外围边线和岛屿轮廓绘出即可。对于缓冲区检索，在外边线所形成的多边形检索后，要再扣除所有岛屿多边形的检索结果。

基于栅格结构也可以作缓冲区分析，通常称为推移或扩散（Spread）。推移或扩散实际上是模拟主体对邻近对象的作用过程，物体在主体的作用下在一阻力表面移动，离主体越远作用力越弱。例如可以将地形、障碍物和空气作为阻力表面，噪声源为主体，用推移或扩散的方法计算噪声离开主体后在阻力表面上的移动，得到一定范围内每个栅格单元的噪声强度。

5．叠加分析

大部分 GIS 软件是以分层的方式组织地理景观，将地理景观按主题分层提取，同一地区的整个数据层集表达了该地区地理景观的内容。每个主题层，可以叫做一个数据层面。数据层面既可以用矢量结构的点、线、面图层文件方式表达，也可以用栅格结构的图层文件格式进行表达。

叠加分析是地理信息系统最常用的提取空间隐含信息的手段之一。该方法源于传统的透明材料叠加，即将来自不同的数据源的图纸绘于透明纸上，在透光桌上将其叠放在一起，然后用笔勾出感兴趣的部分—提取出感兴趣的信息。地理信息系统的叠加分析是将有关主题层组成的数据层面，进行叠加产生一个新数据层面的操作，其结果综合了原来两层或多层要素所具有的属性。叠加分析不仅包含空间关系的比较，还包含属性关系的比较。地理信息系统叠加分析可以分为以下几类：视觉信息叠加、点与多边形叠加、线与多边形叠加、多边形叠加、栅格图层叠加。

5．1 视觉信息叠加

视觉信息叠加是将不同侧面的信息内容叠加显示在结果图件或屏幕上，以便研究者判断其相互空间关系，获得更为丰富的空间信息。地理信息系统中视觉信息叠加包括以下几类：

- 点状图，线状图和面状图之间的叠加显示。
- 面状图区域边界之间或一个面状图与其他专题区域边界之间的叠加。
- 遥感影象与专题地图的叠加。
- 专题地图与数字高程模型（DEM）叠加显示立体专题图。

视觉信息叠加不产生新的数据层面，只是将多层信息复合显示，便于分析。

5．2 点与多边形叠加

点与多边形叠加，实际上是计算多边形对点的包含关系。矢量结构的 GIS 能够通过计算每个点相对于多边形线段的位置，进行点是否在一个多边形中的空间关系判断。

在完成点与多边形的几何关系计算后，还要进行属性信息处理。最简单的方式是将多边形属性信息叠加到其中的点上。当然也可以将点的属性叠加到多边形上，用于标识该多边形，如果有多个点分布在一个多边形内的情形时，则要采用一些特殊规则，如将点的数目或各点属性的总和等信息叠加到多边形上。

通过点与多边形叠加，可以计算出每个多边形类型里有多少个点，不但要区分点是否在多边形内，还要描述在多边形内部的点的属性信息。通常不直接产生新数据层面，只是把属性信息叠加到原图层中，然后通过属性查询间接获得点与多边形叠加的需要信息。例如一个中国政区图（多边形）和一个全国矿产分布图（点），二者经叠加分析后，并且将政区图多边形有关的属性信息加到矿产的属性数据表中，然后通过属性查询，可以查询指定省有多少种矿产，产量有多少；而且可以查询，指定类型的矿产在哪些省里有分布等信息。

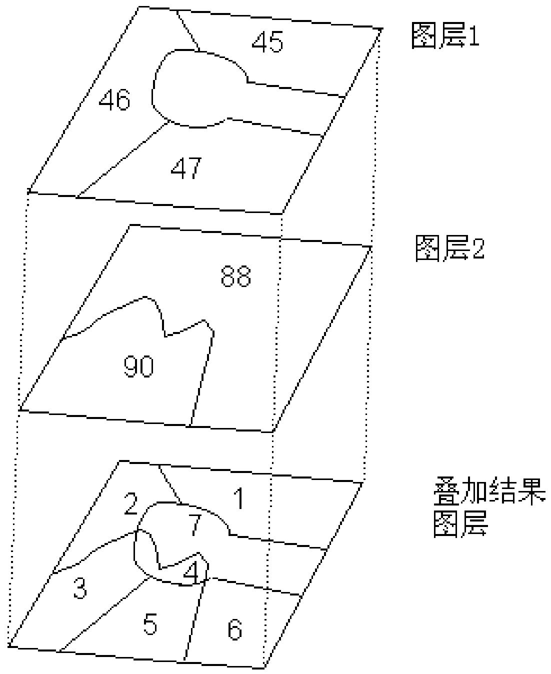
5．3 线与多边形叠加

线与多边形的叠加，是比较线上坐标与多边形坐标的关系，判断线是否落在多边形内。计算过程通常是计算线与多边形的交点，只要相交，就产生一个结点，将原线打断成一条条弧段，并将原线和多边形的属性信息一起赋给新弧段。叠加的结果产生了一个新的数据层面，每条线被它穿过的多边形打断成新弧段图层，同时产生一个相应的属性数据表记录原线和多边形的属性信息。根据叠加的结果可以确定每条弧段落在哪个多边形内，可以查询指定多边形内指定线穿过的长度。如果线状图层为河流，叠加的结果是多边形将穿过它的所有河流打断成弧段，可以查询任意多边形内的河流长度，进而计算它的河流密度等；如果线状图层为道路网，叠加的结果可以得到每个多边形内的道路网密度，内部的交通流量，进入、离开各个多边形的交通量，相邻多边形之间的相互交通量。

5．4 多边形叠加

多边形叠加是 GIS 最常用的功能之一。多边形叠加将两个或多个多边形图层进行叠加产生一个新多边形图层的操作，其结果将原来多边形要素分割成新要素，新要素综合了原来

两层或多层的属性。如图 8-11 所示：



宗地ID	宗地号
45	京-99-01
46	京-99-02
47	京-99-03

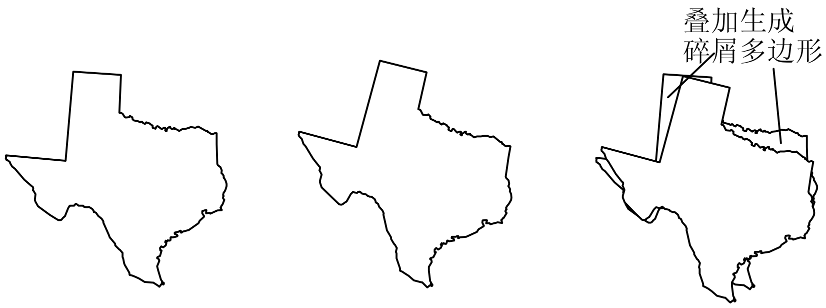
土壤ID	稳定性
88	稳定
90	不稳定

ID	宗地ID	宗地号	土壤ID	稳定性
1	45	京-99-01	88	稳定
2	46	京-99-02	88	稳定
3	46	京-99-02	90	不稳定
4	-	-	90	不稳定
5	47	京-99-03	90	不稳定
6	47	京-99-03	88	稳定
7	-	-	88	稳定

图 8-11：多边形叠加分析

进行多个多边形的叠加运算，在参与运算多边形所构成的属性空间（就图 10 而言，为宗地 ID，宗地号，土壤 ID，稳定性）内，每个结果多边形内部的属性值是一致的，可以称为最小　公共地理单元（Least Common Geographic Unit,LCGU）。

叠加过程可分为几何求交过程和属性分配过程两步。几何求交过程首先求出所有多边形边界线的交点，再根据这些交点重新进行多边形拓扑运算，对新生成的拓扑多边形图层的每个对象赋一多边形唯一标识码，同时生成一个与新多边形对象一一对应的属性表。由于矢量结构的有限精度原因，几何对象不可能完全匹配，叠加结果可能会出现一些碎屑多边形（Silver Polygon），如图 8-12 所示。通常可以设定一模糊容限以消除它。



T1 时刻多边形　　T2 时刻多边形　　多边形叠加结果

图 8-12：多边形叠加产生碎屑多边形

多边形叠加结果通常把一个多边形分割成多个多边形，属性分配过程最典型的方法是将输入图层对象的属性拷贝到新对象的属性表中，或把输入图层对象的标识作为外键，直接关联到输入图层的属性表。这种属性分配方法的理论假设是多边形对象内属性是均质的，将它们分割后，属性不变。也可以结合多种统计方法为新多边形赋属性值。

多边形叠加完成后，根据新图层的属性表可以查询原图层的属性信息，新生成的图层和其它图层一样可以进行各种空间分析和查询操作。

根据叠加结果最后欲保留空间特征的不同要求，一般的 GIS 软件都提供了三种类型的多边形叠加操作，如图 8-13 所示：

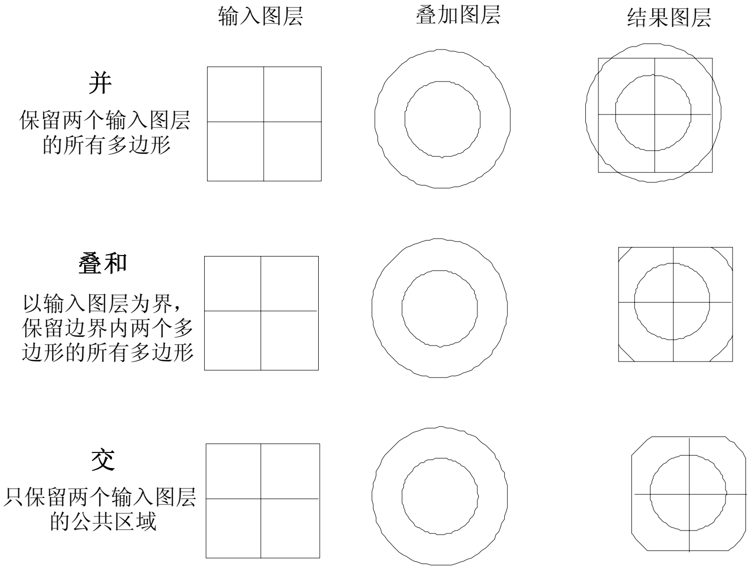


图 8-13：多边形的不同叠加方式

5．5 栅格图层叠加

栅格数据结构空间信息隐含属性信息明显的特点，可以看作是最典型的数据层面，通过数学关系建立不同数据层面之间的联系是 GIS 提供的典型功能。空间模拟尤其需要通过各种各样的方程将不同数据层面进行叠加运算，以揭示某种空间现象或空间过程。例如土壤侵蚀强度与土壤可蚀性，坡度，降雨侵蚀力等因素有关，可以根据多年统计的经验方程，把土壤可蚀性、坡度、降雨侵蚀力作为数据层面输入，通过数学运算得到土壤侵蚀强度分布图。这种作用于不同数据层面上的基于数学运算的叠加运算，在地理信息系统中称为地图代数。地图代数功能有三种不同的类型：

- 基于常数对数据层面进行的代数运算；
- 基于数学变换对数据层面进行的数学变换（指数、对数、三角变换等）；
- 多个数据层面的代数运算（加、减、乘、除、乘方等）和逻辑运算（与、或、非、异或等）。

下面给出一个地图代数的典型例子。有一个森林地区融雪经验模型：

M=（0.19T+0.17D）

式中，M 是融雪速度（厘米/天），T 是空气温度，D 是露点温度。根据此方程，使用该地区的气温和露点温度分布图层，就能计算该地区融雪速率分布图。计算过程是先分别把温度分布图乘以 0.19 和露点温度分布图乘以 0.17，再把得到的结果相加。需要说明的是地图代数在形式和概念上都比较简单，使用起来方便灵活，但把图层作为代数公式的变量进行计算，在实现的技术上难度较大。

栅格图层叠加的另一形式是二值逻辑叠加，常作为栅格结构的数据库查询工具。数据库查询就是查找数据库中已有的信息，例如：基于位置信息查询如已知地点的土地类型，以及基于属性信息的查询如地价最高的位置；比较复杂的查询涉及多种复合条件，如查询所有的面积大于 10 公顷且邻近工业区的全部湿地。这种数据库查询通常分为两步，首先进行再分

类（见第三节）操作，为每个条件创建一个新图层，通常是二值图层，1 代表符合条件，0 表示所有不符合条件。第二步进行二值逻辑叠加操作得到想查询的结果。逻辑操作类型包括与、或、非、异或。

6．网络分析

对地理网络（如交通网络）、城市基础设施网络（如各种网线、电力线、电话线、供排水管线等）进行地理分析和模型化，是地理信息系统中网络分析功能的主要目的。网络分析是运筹学模型中的一个基本模型，它的根本目的是研究、筹划一项网络工程如何安排，并使其运行效果最好，如一定资源的最佳分配，从一地到另一地的运输费用最低等。其基本思想则在于人类活动总是趋于按一定目标选择达到最佳效果的空间位置。这类问题在社会经济活动中不胜枚举，因此在地理信息系统中此类问题的研究具有重要意义。

6．1 网络数据结构

网络数据结构的基本组成部分和属性如下：

1）链（Link）

网络中流动的管线，如街道、河流、水管等，其状态属性包括阻力和需求。

2）结点（Node）[*]

网络中链的结点，如港口、车站、电站等，其状态属性包括阻力和需求等。结点中又有下面几种特殊的类型。

- 障碍（Barrier），禁止网络中链上流动的点。
- 拐点（Turn），出现在网络链中的分割结点上，状态属性有阻力，如拐弯的时间和限制（如在 8：00 到 18：00 不允许左拐）。
- 中心（Center），是接受或分配资源的位置，如水库、商业中心、电站等，其状态属性包括资源容量（如总量），阻力限额（中心到链的最大距离或时间限制）。
- 站点（Stop），在路径选择中资源增减的结点，如库房、车站等，其状态属性有资源需求，如产品数量。

除了基本的组成部分外，有时还要增加一些特殊结构，如邻接点链表用来辅助进行路径分析。

6．2 主要网络分析功能

6．2．1 路径分析

1）静态求最佳路径：在给定每条链上的属性后，求最佳路径。

2）N 条最佳路径分析：确定起点或终点，求代价最小的 N 条路径，因为在实践中最佳路径的选择只是理想情况，由于种种因素而要选择近似最优路径。

3）最短路径或最低耗费路径：确定起点、终点和要经过的中间点、中间连线，求最短

---

[*] 在图数据结构中称为“弧”。

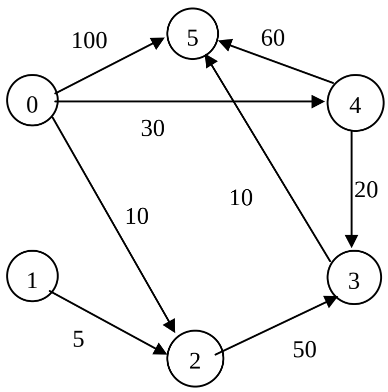
路径或最小耗费路径。

4）动态最佳路径分析：实际网络中权值是随权值关系式变化的，可能还会临时出现一些障碍点，需要动态的计算最佳路径。

6．2．2 计算最短路径的 Dijkstra 算法

为了进行网络最短路径分析，需要将网络转换成有向图。无论是计算最短路径还是最佳路径，其算法都是一致的，不同之处在于有向图中每条弧的权值设置。如果要计算最短路径，则权重设置为两个节点的实际距离；而要计算最佳路径，则可以将权值设置为从起点到终点的时间或费用。Dijkstra 算法可以用于计算从有向图中任意一个节点到其它节点的最短路径。下面是该算法的描述。

1）用带权的邻接矩阵*Cost*来表示带权的*n*个节点的有向图，*Cost[i,j]*表示弧$<v_i,v_j>$的权值，如果从v_i到v_j不连通，则*Cost[i,j]*=∞。图 8-14 表示了一个带权有向图以及其邻接矩阵。



$$\begin{bmatrix} \infty & \infty & 10 & \infty & 30 & 100 \\ \infty & \infty & 5 & \infty & \infty & \infty \\ \infty & \infty & \infty & 50 & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty & 10 \\ \infty & \infty & \infty & 20 & \infty & 60 \\ \infty & \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$

图 8-14：带权的有向图和邻接矩阵

然后，引进一个辅助向量*Dist*，每个分量*Dist[i]*表示从起始点到每个终点v_i的最短路径长度。假定起始点在有向图中的序号为*i0*，并设定该向量的初始值为：

$Dist[i]=Cost[i0,i]$　　　　　　　　　　$v_i \in V$。

令 *S* 为已经找到的从起点出发的最短路径的终点的集合。

2）选择V_j，使得

$Dist[j]=Min\{ Dist[i]|V_i \in V-S\}$　　　　　$v_i \in V$

v_j就是当前求得的一条从v_{i0}出发的最短路径的终点，令

$S=S \cup \{v_j\}$

3）修改从v_{i0}出发到集合*V-S*中任意一顶点v_k的最短路径长度。如果

$Dist[j]+Cost[j,k]<Dist[k]$

则修改 *Dist[k]*为：

$Dist[k]=Dist[j]+Cost[j,k]$

4）重复第 2、3 步操作共*n-1* 次，由此求得从v_{i0}出发的到图上各个顶点的最短路径是依路径长度递增的序列。表 8-1 是图 8-14 根据Dijkstra计算的结果。

表 8-1：用 Dijkstra 计算的结果

终点	从v_0到其它各个节点的最短路径				
v_1	∞	∞	∞	∞	∞ 无
v_2	10 （v_0,v_2）				
v_3	∞	60	50		

		(v_0,v_2,v_3)	(v_0,v_4,v_3)		
v_4	30 (v_0,v_4)	30 (v_0,v_4)			
v_5	100 (v_0,v_5)	100 (v_0,v_5)	90 (v_0,v_4,v_5)	60 (v_0,v_4,v_3,v_5)	
v_j	v_2	v_4	v_3	v_5	

在实际应用中，采用Dijkstra算法计算两点之间的最短路径和求从一点到其它所有点的最短路径所需要的时间是一样的，算法时间复杂度为$O(n^2)$。

6．2．3 资源分配

资源分配网络模型由中心点（分配中心或收集中心）及其属性和网络组成。分配有两种形式，一种是由分配中心向四周分配，另一种是由四周向收集中心分配。资源分配的应用包括消防站点分布和求援区划分、学校选址、垃圾收集站点分布，停水停电对区域的社会、经济影响估计等。

1）负荷设计

负荷设计可用于估计排水系统在暴雨期间是否溢流，输电系统是否超载等。

2）时间和距离估算

时间和距离估算除用于交通时间和交通距离分析外，还可模拟水、电等资源或能量在网络上的距离损耗。

网络分析的具体门类、对象、要求变化非常多，一般的 GIS 软件往往只能提供一些常用的分析方法、或提供描述网络的数据模型和存储信息的数据库。其中最常用的方法是线性阻抗法，即资源在网络上的运输与所受的阻力和距离（或时间）成线性正比关系，在这基础上选择路径，估计负荷，分配资源，计算时间和距离等。对于特殊的、精度要求极高的、非线性阻抗的网络，则需要特殊的算法分析。

7．空间插值

7．1 空间插值的概念和理论

空间插值常用于将离散点的测量数据转换为连续的数据曲面，以便与其它空间现象的分布模式进行比较，它包括了空间内插和外推两种算法。空间内插算法是一种通过已知点的数据推求同一区域其它未知点数据的计算方法；空间外推算法则是通过已知区域的数据，推求其它区域数据的方法。在以下几种情况下必须作空间插值：

1）现有的离散曲面的分辨率，象元大小或方向与所要求的不符，需要重新插值。例如将一个扫描影象（航空像片、遥感影象）从一种分辨率或方向转换到另一种分辨率或方向的影象。

2）现有的连续曲面的数据模型与所需的数据模型不符，需要重新插值。如将一个连续的曲面从一种空间切分方式变为另一种空间切分方式，从 TIN 到栅格、栅格到 TIN 或矢量多边形到栅格。

3）现有的数据不能完全覆盖所要求的区域范围，需要插值。如将离散的采样点数据内插为连续的数据表面。

空间插值的理论假设是空间位置上越靠近的点，越可能具有相似的特征值；而距离越远

的点，其特征值相似的可能性越小。然而，还有另外一种特殊的插值方法——分类，它不考虑不同类别测量值之间的空间联系，只考虑分类意义上的平均值或中值，为同类地物赋属性值。它主要用于地质、土壤、植被或土地利用的等值区域图或专题地图的处理，在“景观单元”或图斑内部是均匀和同质的，通常被赋给一个均一的属性值，变化发生在边界上。

7．2 空间插值的数据源

连续表面空间插值的数据源包括：

- 摄影测量得到的正射航片或卫星影象；
- 卫星或航天飞机的扫描影象；
- 野外测量采样数据，采样点随机分布或有规律的线性分布（沿剖面线或沿等高线）；
- 数字化的多边形图、等值线图；

空间插值的数据通常是复杂空间变化有限的采样点的测量数据，这些已知的测量数据称为“硬数据”。如果采样点数据比较少的情况下，可以根据已知的导致某种空间变化的自然过程或现象的信息机理，辅助进行空间插值，这种已知的信息机理，称为“软信息”。但通常情况下，由于不清楚这种自然过程机理，往往不得不对该问题的属性在空间的变化作一些假设，例如假设采样点之间的数据变化是平滑变化，并假设服从某种分布概率和统计稳定性关系。

采样点的空间位置对空间插值的结果影响很大，理想的情况是在研究区内均匀布点。然而当区域景观大量存在有规律的空间分布模式时，如有规律间隔的数或沟渠，用完全规则的采样网络则显然会得到片面的结果，正是这个原因，统计学家希望通过一些随机的采样来计算无偏的均值和方差。但是完全随机的采样同样存在缺陷，首先随机的采样点的分布位置是不相关的，而规则采样点的分布则只需要一个起点位置，方向和固定大小的间隔，尤其是在复杂的山地和林地里比较容易。其次完全随机采样，会导致采样点的分布不均，一些点的数据密集，另一些点的数据缺少。图 8-15 列出空间采样点分布的几种选择。

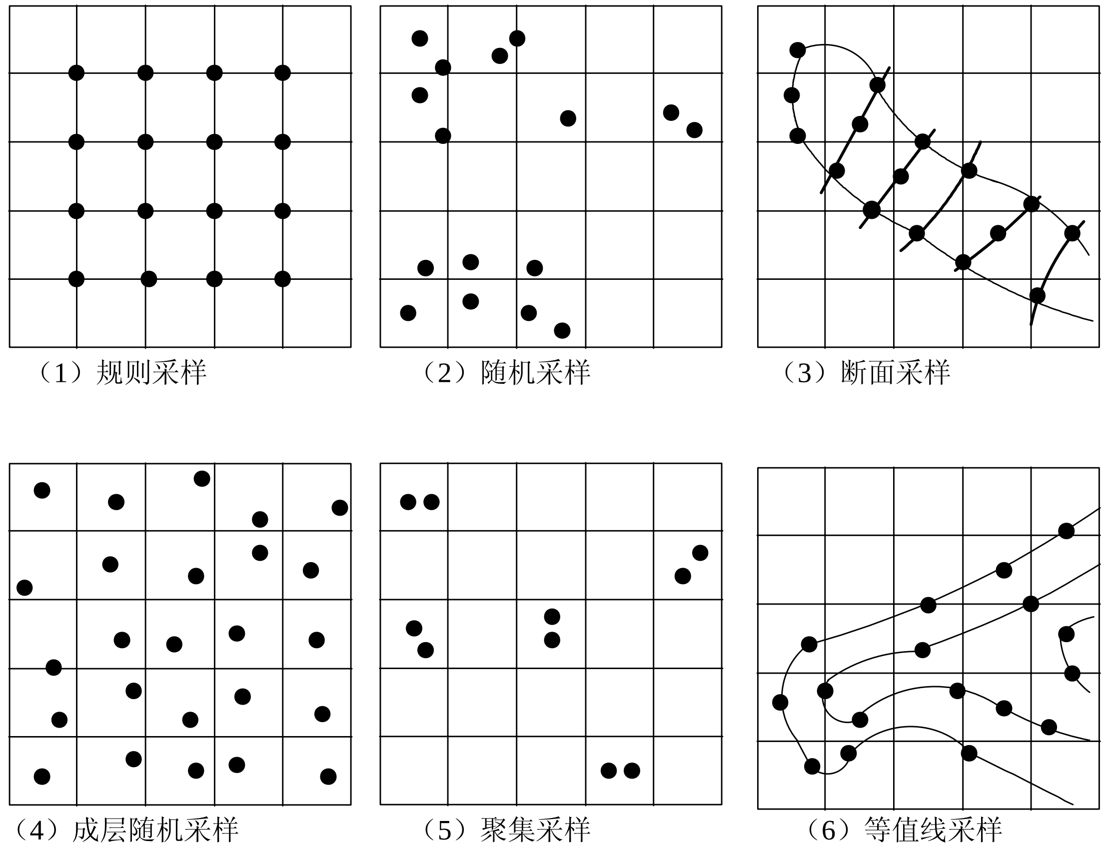


图 8-15：各种不同的采样方式

规则采样和随机采样好的结合方法是成层随机采样，即单个的点随机的分布于规则的格网内。聚集采样可用于分析不同尺度的空间变化。规则断面采样常用于河流、山坡剖面的测量。等值线采样是数字化等高线图插值数字高程模型最常用的方法。

术语：$z(x_i)$是一个数据点的属性值，其中x_i是所有测量点中的一个。$\hat{z}(x_0)$是一个点x_0插值后的数值。如果一种插值方法计算的数据，其中采样点的计算数据等于已知的采样数据，称这种插值方法是精确插值方法；所有的其它插值方法为近似插值方法。统计计算值和测量值之间的差异（绝对值和平方差），$\hat{z}(x_0)-z(x_i)$，是评价不精确插值方法质量常用的指标。

7．3 空间插值方法

空间插值方法可以分为整体插值和局部插值方法两类。整体插值方法用研究区所有采样点的数据进行全区特征拟合；局部插值方法是仅仅用邻近的数据点来估计未知点的值。

整体插值方法通常不直接用于空间插值，而是用来检测不同于总趋势的最大偏离部分，在去除了宏观地物特征后，可用剩余残差来进行局部插值。由于整体插值方法将短尺度的、局部的变化看作随机的和非结构的噪声，从而丢失了这一部分信息。局部插值方法恰好能弥补整体插值方法的缺陷，可用于局部异常值，而且不受插值表面上其它点的内插值影响。

7．3．1 整体插值方法

1）边界内插方法

边界内插方法假设任何重要的变化发生在边界上，边界内的变化是均匀的，同质的，即在各方向都是相同的。这种概念模型经常用于土壤和景观制图，可以通过定义“均质的”土壤单元、景观图斑，来表达其它的土壤、景观特征属性。

边界内插方法最简单的统计模型是标准方差分析（ANOVAR）模型：

$$z(x_0)=\mu+\alpha_k+\varepsilon$$

式中，z是在x_0位置的属性值，μ是总体平均值，α_k是k类平均值与μ的差，ε为类间平均误差（噪声）。

该模型假设每一类别k的属性值是正态分布；每类k的平均值（$\mu+\alpha_k$）由一个独立样品集估计，并假设它们是与空间无关的；类间平均误差ε假设所有类间都是相同的。

评价分类效果的指标是$({\delta_w}^2/{\delta_t}^2)$，${\delta_w}^2$为类间方差，${\delta_t}^2$为总体方差，比值越小分类效果越好。分类效果的显著性检验可以用 F 检验。

实质上，边界内插方法的理论假设是：

- 属性值 z 在“图斑”或景观单元内是随机变化的，不是有规律的；
- 同一类别的所有“图斑”存在同样的类方差（噪声）；
- 所有的属性值都呈正态分布；
- 所有的空间变化发生在边界上，是突变而不是渐变。

在使用边界内插时，应仔细考虑数据源是否符合这些理论假设。

2）趋势面分析

某种地理属性在空间的连续变化，可以用一个平滑的数学平面加以描述。思路是先用已知采样点数据拟合出一个平滑的数学平面方程，再根据该方程计算无测量值的点上的数据。这种只根据采样点的属性数据与地理坐标的关系，进行多元回归分析得到平滑数学平面方程的方法，称为趋势面分析。它的理论假设是地理坐标（*x,y*）是独立变量，属性值 Z 也是独立变量且是正态分布的，同样回归误差也是与位置无关的独立变量。

多项式回归分析是描述长距离渐变特征的最简单方法。多项式回归的基本思想是用多项式表示线、面，按最小二乘法原理对数据点进行拟合。线或面多项式的选择取决于数据是一维的还是二维的。

用一个简单的示例来说明，地理或环境调查中特征值z沿一个断面在x_1，$x_2 \ldots x_n$处采样，若z值随x值增加而线性增大，则该特征值的长期变化可以用下面一个回归方程进行计算：

$$z(x)=b_0+b_1x+\varepsilon$$

其中，b_0，b_1为回归系数，ε为独立于x的正态分布残差（噪声）。

然而许多情况下，不是以线性函数，而是以更为复杂的方式变化，则需用二次多项式进行拟合：

$$z(x)=b_0+b_1x+b_2x^2+\varepsilon$$

对于二维的情况，XY 坐标的多元回归分析得到的曲面多项式，形式如下：

$$f\{(x,y)\}=\sum_{r+s\le p}\left(b_{rs}\cdot x^r\cdot y^s\right)$$

前三种形式分别是：

b_0 平面

$b_0 + b_1 x + b_2 y$ 斜平面

$b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 xy + b_5 y^2$ 二次曲面

其中，p 是趋势面方程的次数。

$P = (p+1)(p+2)/2$ 是趋势面多项式正常情况下的最少项数个数。零次多项式是平面，有 1 个项数；一次多项式是斜平面，有 3 个项数；二次曲面有 6 个项数，三次趋势面有 10 个项数。

计算系数b_i是一个标准的多元回归问题。趋势面分析的优点是非常容易理解，至少是在计算方面。另外大多数情况下可用低次多项式进行拟合，但给复杂的多项式赋与明确的物理意义比较困难。

趋势面是个平滑函数，很难正好通过原始数据点，除非是数据点少且趋势面次数高才能是曲面正好通过原始数据点，所以趋势面分析是一个近似插值方法。实际上趋势面最有成效的应用是揭示区域中不同于总趋势的最大偏离部分，所以趋势面分析的主要用途是，在使用某种局部插值方法之前，可用趋势面分析从数据中去掉一些宏观特征，不直接用它进行空间插值。

趋势面拟合程度的检验，同多元回归分析一样，可用 F 分布进行检验，其检验统计量为：

$$F = \frac{U/p}{Q/(n-p-1)}$$

其中，U为回归平方和，Q为残差平方和（剩余平方和），p为多项式项数（不包括常数项b_0），n为使用数据点数目。当$F>F_a$时，趋势面显著，否则不显著。

3）变换函数插值

根据一个或多个空间参量的经验方程进行整体空间插值，也是经常使用的空间插值方法，这种经验方程称为变换函数。下面以一个研究实例进行说明。

冲积平原的土壤重金属污染与几个重要因子有关，其中距污染源（河流）的距离，和高程两个因子最重要。一般情况，携带重金属的粗粒泥沙沉积在河滩上，携带重金属的细粒泥沙沉淀在低洼的在洪水期容易被淹没的地方，而那些洪水频率低的地方，由于携带重金属污染泥沙颗粒比较少，受到污染轻。由于距河流的距离和高程是比较容易得到的空间变量，可以用各种重金属含量与它们的经验方程进行空间插值，以改进对重金属污染的预测。本例回归方程的形式如下：

$$z(x) = b_0 + b_1 p_1 + b_2 p_2 + \varepsilon$$

式中是z(x)某种重金属含量（ppm），$b_0 \ldots b_n$是回归系数，$p_1 \ldots p_n$是独立空间变量，本例p_1是距河流的距离因子，p_2是高程因子。

这种回归模型通常叫做转换函数，大多数 GIS 软件都可以计算。转换函数可以应用于其他独立变量，如温度、高程、降雨量和距海、植被的距离关系可以组合为一个超剩含水量的函数。地理位置及其属性可以尽可能多的信息组合成需要的回归模型，然后进行空间插值。但应该注意的一点是，必须清楚回归模型的物理意义。还要指出的是所有的回归转换函数都属于近似的空间插值。

整体插值方法通常使用方差分析和回归方程等标准的统计方法，计算比较简单。其它的许多方法也可用于整体空间插值，如傅立叶级数和小波变换，特别是遥感影象分析方面 ，但它们需要的数据量大。

7．3．2 局部插值方法

局部插值方法只使用邻近的数据点来估计未知点的值，包括几个步骤：

1）定义一个邻域或搜索范围；

2）搜索落在此邻域范围的数据点；

3）选择表达这有限个点的空间变化的数学函数；

4）为落在规则格网单元上的数据点赋值。重复这个步骤直到格网上的所有点赋值完毕。

使用局部插值方法需要注意的几个方面是：所使用的插值函数；邻域的大小、形状和方向；数据点的个数；数据点的分布方式是规则的还是不规则的。

1）最近邻点法：泰森多边形方法

泰森多边形（Thiessen，又叫 Dirichlet 或 Voronoi 多边形）采用了一种极端的边界内插方法，只用最近的单个点进行区域插值。泰森多边形按数据点位置将区域分割成子区域，每个子区域包含一个数据点，各子区域到其内数据点的距离小于任何到其它数据点的距离，并用其内数据点进行赋值。连接所有数据点的连线形成 Delaunay 三角形，与不规则三角网 TIN 具有相同的拓扑结构。

GIS 和地理分析中经常采用泰森多边形进行快速的赋值，实际上泰森多边形的一个隐含的假设是任何地点的气象数据均使用距它最近的气象站的数据。而实际上，除非是有足够多的气象站，否则这个假设是不恰当的，因为降水、气压、温度等现象是连续变化的，用泰森多边形插值方法得到的结果图变化只发生在边界上，在边界内都是均质的和无变化的。

2）移动平均插值方法：距离倒数插值

距离倒数插值方法综合了泰森多边形的邻近点方法和趋势面分析的渐变方法的长处，它假设未知点x_0处属性值是在局部邻域内中所有数据点的距离加权平均值。距离倒数插值方法是加权移动平均方法的一种。加权移动平均方法的计算公式如下：

$$\hat{z}(x_0)=\sum_{i=1}^{n}\lambda_i\cdot z(x_i) \qquad \sum_{i=1}^{n}\lambda_i=1$$

式中，权重系数由函数$\phi(d(x,x_i))$计算，要求当$d\to 0$时$\phi(d)\to 1$，一般取倒数或负指数形式d^{-r},e^{-d},e^{-d^2}。其中$\phi(d(x,x_i))$最常见的形式是距离倒数加权函数，形式如下：

$$\hat{z}(x_j)=\sum_{i=1}^{n}z(x_i)\cdot d_{ij}^{-r}\Big/\sum_{i=1}^{n}d_{ij}^{-r}$$

式中，x_j为未知点，x_i为已知数据点。

加权移动平均公式最简单的形式称为线性插值，公式如下：

$$\hat{z}(x_0)=\frac{1}{n}\sum_{i=1}^{n}z(x_i)$$

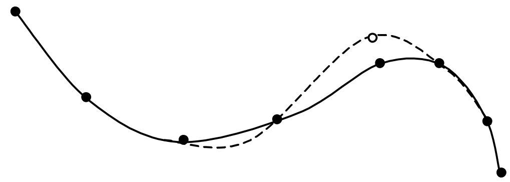
距离倒数插值方法是 GIS 软件根据点数据生成栅格图层的最常见方法。距离倒数法计算值易受数据点集群的影响，计算结果经常出现一种孤立点数据明显高于周围数据点的“鸭

蛋”分布模式，可以在插值过程中通过动态修改搜索准则进行一定程度的改进。

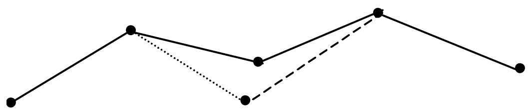
3）样条函数插值方法

在计算机用于曲线与数据点拟合以前，绘图员是使用一种灵活的曲线规逐段的拟合出平滑的曲线。这种灵活的曲线规绘出的分段曲线称为样条。与样条匹配的那些数据点称为桩点，绘制曲线时桩点控制曲线的位置。曲线规绘出的曲线在数学上用分段的三次多项式函数来描述这种曲线，其连接处有连续的一阶和二阶连续导数。

样条函数是数学上与灵活曲线规对等的一个数学等式，是一个分段函数，进行一次拟合只有与少数点拟合，同时保证曲线段连接处连续。这就意味着样条函数可以修改少数数据点配准而不必重新计算整条曲线，趋势面分析方法做不到这一点，（图 16）。



(a)



(b)

图 8-16：样条函数的局部特征

（a：当二次样条曲线的一个点位置变化时，只需要重新计算四段曲线；

b：而一次样条曲线的一个点位置变化时，只需要重新计算两段曲线）

一般的分段多项式 $p(x)$定义为：

$p(x)=p_i(x) \qquad x_i < x < x_{i+1} \qquad (i=1, 2, 3\ldots, k-1)$

$$p_i^{(j)}(x_i) = p_{i+1}^{(j)}(x_i) \quad (j=0, 1, 2, ..., r-1; i=1, 2, ... , k-1)$$

$x_1, ..., x_{k-1}$将区间x_0，x_k分成k个子区间，这些分割点称“断点”，曲线上具有这些x值的点称为“节”。函数$p_i(x)$为小于等于m次的多项式。r项用来表示样条函数的约束条件：

$r=0$，无约束；

$r=1$，函数连续且对它的导数无任何约束；

$r=m-1$，区间$[x_0 x_k]$可用一个多项式表示；

r=m，约束条件最多。

m=1,2,3 时的样条分别为一次、二次、三次样条函数，其导数分别是 0 阶、1 阶、2 阶导数，二次样条函数的每个节点处必须有一阶连续导数，三次样条函数的每个节点初必须有二阶连续导数。$r=m$ 的简单样条只有 $k+m$ 个自由度，$r=m=3$ 有着特殊的意义，因为它是三次多项式，该函数首次被人们称为样条函数。术语“三次样条”用于三维情况，此时进行曲面内插，而不是曲线内插。

由于离散子区间的范围较宽，可能是一条数字化的曲线，在这个范围内计算简单样条会引起一定的数学问题，因此在实际应用中都用 *B* 样条──一种特殊的样条函数。*B* 样条是感兴趣区间以外均为零的其它样条的和，因此可按简单的方法用低次多项式进行局部拟合。

B 样条经常用于数字化的线划在显示之前进行平滑处理，例如土壤、地质图上的各种边界，传统的制图总希望绘出较平滑的曲线。但是用 *B* 样条做多边形边界平滑也存在一些问题，特别是多边形面积和周长的计算，结果会与平滑前的不同。

综上所述，样条函数是分段函数，每次只用少量数据点，故插值速度快。样条函数与趋势面分析和移动平均方法相比，它保留了局部的变化特征。线性和曲面样条函数都在视觉上上得到了令人满意的结果。样条函数的一些缺点是：样条内插的误差不能直接估算，同时在实践中要解决的问题是样条块的定义以及如何在三维空间中将这些“块”拼成复杂曲面，又不引入原始曲面中所没有的异常现象等问题。

4）空间自协方差最佳插值方法：克里金插值

前面介绍的几个插值方法对影响插值效果的一些敏感性问题仍没有得到很好的解决，例如趋势面分析的控制参数和距离倒数插值方法的权重对结果影响很大，这些问题包括：

- 需要计算平均值数据点的数目；
- 搜索数据点的邻域大小、方向和形状如何确定；
- 有没有比计算简单距离函数更好的估计权重系数的方法；
- 与插值有关的误差问题。

为解决这些问题，法国地理数学学家 Georges Matheron 和南非矿山工程师 D.G.Krige 研究了一种优化插值方法，用于矿山勘探。这个方法被广泛地应用于地下水模拟、土壤制图等领域，成为 GIS 软件地理统计插值的重要组成部分。这种方法充分吸收了地理统计的思想，认为任何在空间连续性变化的属性是非常不规则的，不能用简单的平滑数学函数进行模拟，可以用随机表面给予较恰当的描述。这种连续性变化的空间属性称为“区域性变量”，可以描述象气压、高程及其它连续性变化的描述指标变量。这种应用地理统计方法进行空间插值的方法，被称为克里金（Kriging）插值。 地理统计方法为空间插值提供了一种优化策略，即在插值过程中根据某种优化准则函数动态的决定变量的数值。Matheron ，Krige 等人研究的插值方法着重于权重系数的确定，从而使内插函数处于最佳状态，即对给定点上的变量值提供最好的线性无偏估计。

克里金插值方法的区域性变量理论假设任何变量的空间变化都可以表示为下述三个主要成分的和（图 8-17）：

1）与恒定均值或趋势有关的结构性成分；

2）与空间变化有关的随机变量，即区域性变量；

3）与空间无关的随机噪声项或剩余误差项。

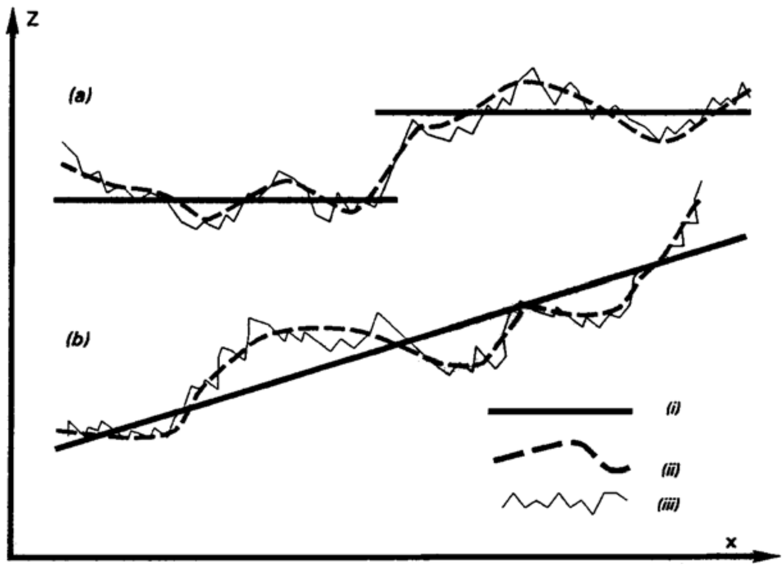


图 8-17：区域变量理论将复杂的空间变化分为三个部分

(i)地形的平均特性；(ii)空间相关的不规则变化；(iii)随机的、局部的变化

另 x 为一维、二维或三维空间中的某一个位置，变量 z 在 x 处的值可又下式计算：

$$z(x)=m(x)+\varepsilon^{'}(x)+\varepsilon^{''}$$

式中，m(x)是描述 z(x)的结构性成分的确定性函数；$\varepsilon^{'}(x)$ 是与空间变化有关的随机变化项，即区域性变量；$\varepsilon^{''}(x)$ 是剩余误差项，空间上具有零平均值、σ^2 与空间无关的高斯噪声项。

克里金方法的第一步是确定适当的*m(x)*函数，最简单的情况是*m(x)*等于采样区的平均值，距离矢量*h*分离的两点<u>*x*</u>，*x+h*之间的数学期望等于零：

$$E[z(x)-z(x+h)]=0$$

式中 *z(x), z(x+h)* 是随机变量 *z* 在 *x, x+h* 处的值，同时还假设两点之间的方差只与距离 *h* 有关，于是：

$$E\left[(z(x)-z(x+h))^2\right]=E\left[(\varepsilon^{'}(x)-\varepsilon^{'}(x+h))^2\right]=2\gamma(h)$$

式中，$\gamma(h)$是一种函数，称为半方差函数。

区域性变量理论的两个内在假设条件是差异的稳定性和可变性，一旦结构性成分确定后，剩余的差异变化属于同质变化，不同位置之间的差异仅是距离的函数。这样，区域性变量计算公式可以写成下式的形式：

$$z(x)=m(x)+\gamma(h)+\varepsilon^{''}$$

半方差的估算公式如下：

$$\hat{\gamma}(h)=\frac{1}{2n}\sum_{i=1}^{n}[z(x_i)-z(x_i+h)]^2$$

式中，*n* 为距离为 *h* 的采样点对的数目（*n* 对点），采样间隔 *h* 也叫延迟。对应于 *h* 的 $\hat{\gamma}(h)$ 的图被称为“半方差图”。图 8-18 表示一个典型的半方差图。

半方差是定量描述区域性变化的第一步，它为空间插值、优化采样方案提供了有益的信息。为了求得半方差图，必须先得到拟合半方差的理论模型，在半方差理论模型中：

1）延迟 *h* 的值较大的部分曲线呈水平方向。曲线的水平部分成为“梁（Sill）”。说明在延迟的这个范围内数据点没有空间相关性，因为所有的方差不随距离增减而变化。

2）曲线从$\gamma(h)$的低值升到梁为止的延迟范围，称为“变程（Range）”。变程是半方差图最重要的部分，因为它描述了与空间有关的差异怎样随距离变化的。在变程范围内距离越近的点具有更相近的特征。变程给移动加权平均方法提供了一个确定窗口大小的方法。很显然，数据点和未知点之间的距离大于变程范围，表明该数据点与未知点距离太远，对插值没有作用。

3）图中的拟合模型没有通过原点，而是在$\gamma(h)$的正方向与坐标轴相截。

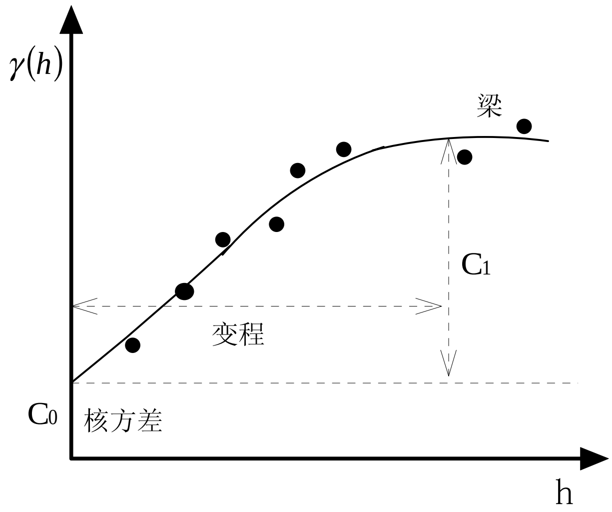


图 8-18：半方差图

按半方差计算公式，当 h=0 时，$\hat{\lambda}(h)$必须为零。模型中出现的正值$\gamma(h)$ $h \to 0$是剩余误差ε''的估计值，它是与空间无关的噪声。ε''称为“核(Nugget)”方差，是观测误差的和距离间隔很小的情况下的空间变化的组合。

当存在明显的变程和梁，同时核方差也很重要但数值不太大的情况下，可用球面模型进行半方差拟合（图 8-19-1）。公式是：

$$\gamma(h) = c_0 + c_1\left[\frac{3h}{2a} - (h/a)^3/2\right] \qquad 0 < h < a$$

$$\gamma(h) = c_0 + c_1 \qquad h >= a$$

$$\gamma(0) = 0 \qquad h=0$$

式中，a是变程，h是延迟，c_0+c_1为梁。一般情况下用球面模型拟合效果比较理想。

如果存在明显的核方差和梁，而没有渐变的变程，则可用指数模型（图 8-19-2）进行拟合：

$$\gamma(h) = c_0 + c_1[1 - \exp(-h/a)]$$

如果核方差相对于与空间变化有关的随机变化$\varepsilon''(x)$很小的情况下，最好使用比较弯曲的曲线，如高斯曲线：（图 8-19-4）

$$\gamma(h) = c_0 + c_1\left[1 - \exp(-h/a)^2\right]$$

如果空间变化随变程渐变，但没有梁，则可用线性模型（图 8-19-3）进行拟合：

$$\gamma(h) = c_0 + bh$$

式中，b 为线的斜率。当变程的大小远超过人们希望的插值范围时，也用线性模型。

前面的讨论都假设地表特征在各个方向都是相同的，然而许多情况下空间变化中的$\varepsilon'(x)$都具有明显的方向性，这是就要用不同参数的模型来拟合半方差图。

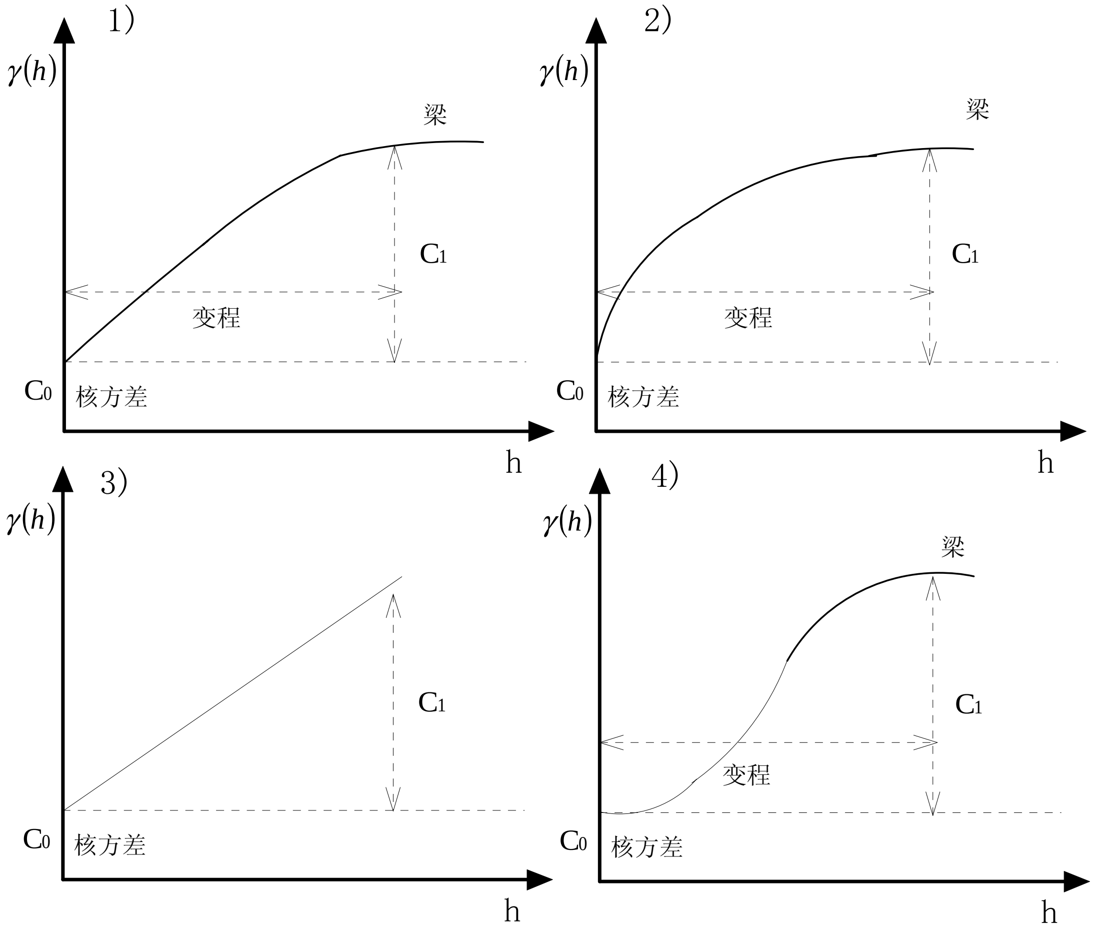


图 8-19：各种不同的变方差图

拟合后的半方差图重要的用途是确定局部内插需要的权重因子 λ_i。确定 λ_i 的过程与加权移动插值方法类似，但不是按一种固定的函数计算 λ_i，而是按采样点数据的半方差图的统计分析原理计算。即：

$$\hat{z}(x_0)=\sum_{i=1}^{n}\lambda_i\cdot z(x_i) \qquad\qquad \sum_{i=1}^{n}\lambda_i=1$$

权重 λ_i 的选择应使 $\hat{z}(x_0)$ 是无偏估计，且估计的方差 σ_e^2 小于观测值的其它线性组合产生的方差。

$\hat{z}(x_0)$ 的最小方差为：

$$\hat{\sigma}_e^2=\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_0)+\Phi$$

只有下式成立时，才可获得 $\hat{z}(x_0)$ 的最小方差：

$$\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_j)+\Phi=\gamma(x_j,x_0)$$

式中，$\gamma(x_i, x_j)$是z在采样点x_i, x_j之间的半方差；$\gamma(x_j, x_0)$是采样点x_j和未知点x_0之间的半方差，这两个量均可从已拟合模型的半方差图上得到。量Φ计算最小方差需要的拉格朗日算子。

这个方法叫常规克里金插值。它是一个精确插值模型，内插值或最佳局部均值与数据点上的值一致。制图中常用比采样间隔更细的规则格网进行插值，内插值又可用前边提到的方法转换成等值线图。与此类似，估计的误差σ_e^2，又叫克里金方差，可以用来制图以反映在整个研究区内插值结果的可靠性。

一个克里金插值的示例：（图 **8-20**）

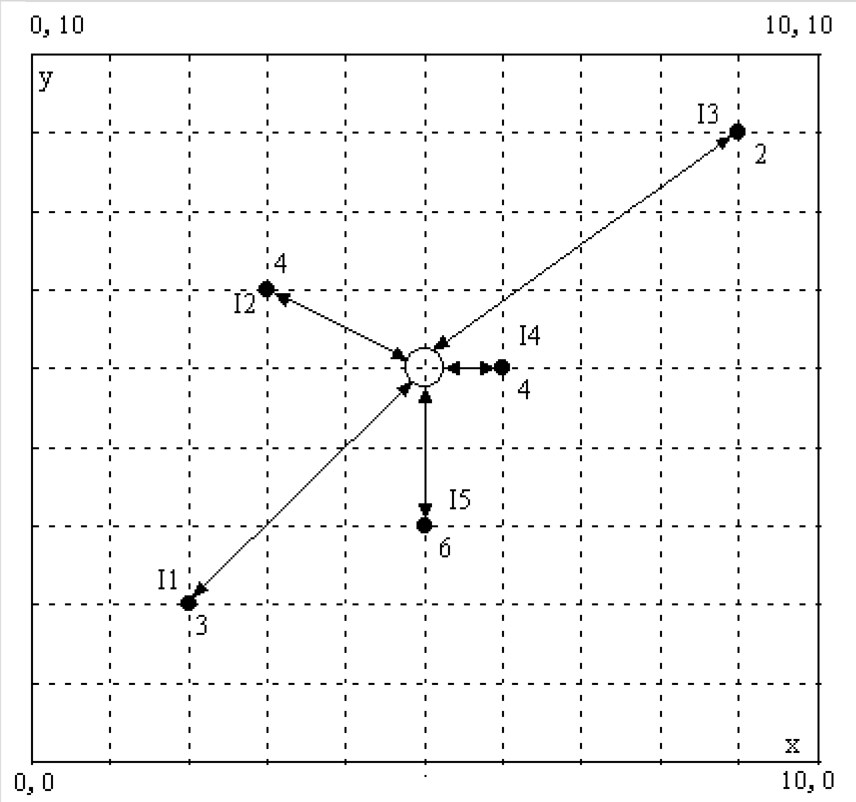


图 8-20：克里金插值的示例

计算图中(x=5,y = 5)处未知点I0 的克里金方法的权重，使用球面模型对半方差进行拟和，参数c_0 = 2.5 ，c_1 =7.5；变程 a= 10.0。这 5 个采样点的数据分别是：

I	x	y	z
I1	2	2	3
I2	3	7	4
I3	9	9	2
I4	6	5	4
I5	5	3	6

解法过程用矩阵的表示如下：

$$\mathrm{A}^{-1}b = \begin{bmatrix} \lambda \\ \Phi \end{bmatrix}$$

其中，A 为数据点之间的半方差矩阵，b 是每个数据点与未知点之间的半方差向量，λ为要计算的权重系数向量，Φ为解方程的拉格朗日算子。

首先计算这 5 个数据点之间的距离矩阵：

I	*1*	*2*	*3*	*4*	*5*
I1	*0.0*	*5.099*	*9.899*	*5.000*	*3.162*
I2	*5.099*	*0.0*	*6.325*	*3.606*	*4.472*
I3	*9.899*	*6.325*	*0.0*	*5.0*	*7.211*
I4	*5.0*	*3.606*	*5.0*	*0.0*	*2.236*
I5	*3.162*	*4.472*	*7.211*	*2.236*	*0.0*

和它们与未知点之间的距离向量：

I	*I0*
I1	*4.243*
I2	*2.808*
I3	*5.657*
I4	*1.0*
I5	*2.0*

将上述数值带入球面模型，得到相应的半方差（矩阵 A，b）：

A=

I	*1*	*2*	*3*	*4*	*5*	*6*
1	*2.500*	*7.739*	*9.999*	*7.656*	*5.939*	*1.000*
2	*7.739*	*2.500*	*8.677*	*6.381*	*7.196*	*1.000*
3	*9.999*	*8.677*	*2.500*	*7.656*	*9.206*	*1.000*
4	*7.656*	*6.381*	*7.656*	*2.500*	*4.936*	*1.000*
5	*5.939*	*7.196*	*9.206*	*4.936*	*2.500*	*1.000*
6	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	*0.000*

b =

I	*0*
1	*7.151*
2	*5.597*
3	*8.815*
4	*3.621*
5	*4.720*
6	*1.000*

注意额外的第 6 行和第 6 列，是为了保证权重之和为 1。

计算 A 的逆矩阵，得：

A^{-1}=

I	*1*	*2*	*3*	*4*	*5*	*6*
1	*-0.172*	*0.050*	*0.022*	*-0.026*	*0.126*	*0.273*
2	*0.050*	*-0.167*	*0.032*	*0.077*	*0.007*	*0.207*
3	*0.022*	*0.032*	*-0.111*	*0.066*	*-0.010*	*0.357*
4	*-0.026*	*0.077*	*0.066*	*-0.307*	*0.190*	*0.030*
5	*0.126*	*0.007*	*-0.010*	*0.190*	*-0.313*	*0.134*
6	*0.273*	*0.207*	*0.357*	*0.003*	*0.134*	*-6.873*

于是，权重 λ 为：

I	权重系数		距离
1	0.0175		4.423
2	0.2281		2.828
3	-0.0891	Σ = 1	5.657
4	0.6437		1.000
5	0.1998		2.000
6	0.1182	φ	

得未知点插值后得数值为：

$$z(I0)=0.0175*3+0.2281*4-0.0891*2+0.6437*4+0.1998*6$$

$$=4.566$$

估计方差为：

$$\sigma_e^2=[0.0175*7.151+0.2281*5.597-0.0891*8.815$$

$$+0.6437*3.621+.1998*4.720]+\phi$$

$$=3.890+0.1182$$

$$=4.008$$

克里金插值方法的目的是提供确定权重系数最优的方法和并能描述误差信息。由于克里金点模型（常规克里金模型）的内插值与原始样本的容量有关，当样本少的情况下，采用简单的点常规克里金插值的内插结果图会出现明显的凹凸现象。可以通过修改克里金方程以估计子块 B 内的平均值来克服这一缺点。该方法叫块克里金插值，该方法对估算给定面积试验小区的平均值或对给定格网大小的规则格网进行插值比较适用。

子块 B 内 z 的均值为：

$$z(B)=\int_B \frac{z(x)dx}{S_B}$$

式中，S_B为子块B的面积。z(x)仍用相同的估计公式：

$$\hat{z}(x_0)=\sum_{i=1}^{n}\lambda_i\cdot z(x_i) \qquad \sum_{i=1}^{n}\lambda_i=1$$

最小方差仍为：

$$\hat{\sigma}_e^2=\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_0)+\Phi$$

但成立条件则变为：

$$\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_j)+\Phi=\bar{\gamma}(x_j,B)$$

块克里金插值估算的方差结果常常小于点克里金插值，所以生成的平滑插值表面不会发生点模型的凹凸现象。

8. 空间统计分类分析

多变量统计分析主要用于数据分类和综合评价。数据分类方法是地理信息系统重要的组成部分。一般说地理信息系统存储的数据具有原始性质，用户可以根据不同的实用目的，进行提取和分析，特别是对于观测和取样数据，随着采用分类和内插方法的不同，得到的结果

有很大的差异。因此，在大多数情况下，首先是将大量未经分类的数据输入信息系统数据库，然后要求用户建立具体的分类算法，以获得所需要的信息。

综合评价模型是区划和规划的基础。从人类认识的角度来看有精确的和模糊的两种类型，因为绝大多数地理现象难以用精确的定量关系划分和表示，因此模糊的模型更为实用，结果也往往更接近实际。综合评价一般经过四个过程：

1）评价因子的选择与简化；

2）多因子重要性指标（权重）的确定；

3）因子内各类别对评价目标的隶属度确定；

4）选用某种方法进行多因子综合。

分类和评价的问题通常涉及大量的相互关联的地理因素，主成分分析方法可以从统计意义上将各影响要素的信息压缩到若干合成因子上，从而使模型大大地简化；因子权重的确定是建立评价模型的重要步骤，权重正确与否极大地影响评价模型的正确性，而通常的因子权重确定依赖较多的主观判断，层次分析法是综合众人意见，科学地确定各影响因子权重的简单而有效的数学手段。隶属度反映因子内各类别对评价目标的不同影响，依据不同因子的变化情况确定，常采用分段线性函数或其它高次函数形式计算。常用的分类和综合的方法包括聚类分析和判别分析两大类。聚类分析可根据地理实体之间影响要素的相似程度，采用某种与权重和隶属度有关的距离指标，将评价区域划分若干类别；判别分析类似于遥感图像处理的分类方法，即根据各要素的权重和隶属度，采用一定的评价标准将各地理实体判归最可能的评价等级或以某个数据值所示的等级序列上；分类定级是评价的最后一步，将聚类的结果根据实际情况进行合并，并确定合并后每一类的评价等级，对于判别分析的结果序列采用等间距或不等间距的标准划分为最后的评价等级。

下面简要介绍分类评价中常用的几种数学方法。

8．1 主成分分析（Principal Component Analysis，PCA）

地理问题往往涉及大量相互关联的自然和社会要素，众多的要素常常给模型的构造带来很大困难，同时也增加了运算的复杂性。为使用户易于理解和解决现有存储容量不足的问题，有必要减少某些数据而保留最必要的信息。由于地理变量中许多变量通常都是相互关联的，就有可能按这些关联关系进行数学处理达到简化数据的目的。主成分分析是通过数理统计分析，求得各要素间线性关系的实质上有意义的表达式，将众多要素的信息压缩表达为若干具有代表性的合成变量，这就克服了变量选择时的冗余和相关，然后选择信息最丰富的少数因子进行各种聚类分析，构造应用模型。

设有n个样本，户个变量。将原始数据转换成一组新的特征值——主成分，主成分是原变量的线性组合且具有正交特征。即将x_1，x_2，…，x_p综合成m（$m<p$）个指标z_1，z_2，…，z_m，即：

$$z_1=l_{11}*x_1+l_{12}*x_2+\ldots+l_{1p}*x_p$$
$$z_2=l_{21}*x_1+l_{22}*x_2+\ldots+l_{2p}*x_p$$
……
$$z_m=l_{m1}*x_1+l_{m2}*x_2+\ldots+l_{mp}*x_p$$

这样决定的综合指标z_1，z_2，…，z_m分别称做原指标的第一，第二，…，第m主成分。其中z_1在总方差中占的比例最大，其余主成分z_2，z_3，…，z_m的方差依次递减。在实际工作中常挑选前几个方差比例最大的主成分，这样既减少了指标的数目，又抓住了主要矛盾，简化了指标之间的关系。

从几何上看，确定主成分的问题，就是找p维空间中椭球体的主轴问题，就是得到即将

x_1，x_2，…，x_p的相关矩阵中m个较大特征值所对应的特征向量，通常用雅可比（Jacobi）法计算特征值和特征向量。

很显然，主成分分析这一数据分析技术是把数据减少到易于管理的程度，也是将复杂数据变成简单类别便于存储和管理的有力工具。

8．2 层次分析法

层次分析（Analytic Hierarchy Process,AHP）法是系统分析的数学工具之一，它把人的思维过程层次化、数量化，并用数学方法为分析、决策、预报或控制提供定量的依据。事实上这是一种定性和定量分析相结合的方法。在模型涉及大量相互关联、相互制约的复杂因素的情况下，各因素对问题的分析有着不同的重要性，决定它们对目标重要性的序列，对建立模型十分重要。

AHP 方法把相互关联的要素按隶属关系分为若干层次，请有经验的专家对各层次各因素的相对重要性给出定量指标，利用数学方法综合专家意见给出各层次各要素的相对重要性权值，作为综合分析的基础。

8．3 系统聚类分析

系统聚类是根据多种地学要素对地理实体进行划分类别的方法，对不同的要素划分类别往往反映不同目标的等级序列，如土地分等定级、水土流失强度分级等。

系统聚类的步骤一般是根据实体间的相似程度，逐步合并若干类别，其相似程度由距离或者相似系数定义。进行类别合并的准则是使得类间差异最大，而类内差异最小。

8．4 判别分析

判别分析与聚类分析同属分类问题，所不同的是，判别分析是预先根据理论与实践确定等级序列的因子标准，再将待分析的地理实体安排到序列的合理位置上的方法，对于诸如水土流失评价、土地适宜性评价等有一定理论根据的分类系统定级问题比较适用。

判别分析依其判别类型的多少与方法的不同，可分为两类判别、多类判别和逐步判别等。

通常在两类判别分析中，要求根据已知的地理特征值进行线性组合，构成一个线性判别函数 Y，即：

$$Y= c_1*x_1+c_2*x_2+\ldots+c_m*x_p$$

式中，c_k（$k=1，2，\ldots，m$）为判别系数，它可反映各要素或特征值作用方向、分辨能力和贡献率的大小。只要确定了c_k，判别函数Y也就确定了。在确定判别函数后，根据每个样本计算判别函数数值，可以将其归并到相应的类别中。常用的判别分析有距离判别法、Bayes 最小风险判别、费歇准则判别等等。

南北美洲三角的西端，都有北南走向的山脉。山脉面积大约占总面积的三分之一，其余的三分之二在东部，是一块大平原。平原被两座低矮的山脉把它与大海隔离开来。在北美洲，是拉布拉多山和阿巴拉契亚山；南美洲则是圭亚那山脉和巴西高原。

房龙

横看成岭侧成峰，远近高低各不同。

苏轼

第九章 数字地形模型（DTM）与地形分析

导读：DEM 和 DTM 主要用于描述地面起伏状况，可以用于提取各种地形参数，如坡度、坡向、粗糙度等，并进行通视分析、流域结构生成等应用分析。因此，DEM 在各个领域中被广泛使用。

DEM 可以有多种表达方法，包括网格、等高线、三角网等，本章同时介绍了这些表达方法之间的相互转换算法，如由三角网生成等高线，网格 DEM 生成三角网等等。

1．概述

数字地形模型（DTM, Digital Terrain Model）最初是为了高速公路的自动设计提出来的（Miller，1956）。此后，它被用于各种线路选线（铁路、公路、输电线）的设计以及各种工程的面积、体积、坡度计算，任意两点间的通视判断及任意断面图绘制。在测绘中被用于绘制等高线、坡度坡向图、立体透视图，制作正射影像图以及地图的修测。在遥感应用中可作为分类的辅助数据。它还是地理信息系统的基础数据，可用于土地利用现状的分析、合理规划及洪水险情预报等。在军事上可用于导航及导弹制导、作战电子沙盘等。对 DTM 的研究包括 DTM 的精度问题、地形分类、数据采集、DTM 的粗差探测、质量控制、数据压缩、DTM 应用以及不规则三角网 DTM 的建立与应用等。

1．1 DTM 和 DEM

从数学的角度，高程模型是高程 Z 关于平面坐标 X，Y 两个自变量的连续函数，数字高程模型（DEM）只是它的一个有限的离散表示。高程模型最常见的表达是相对于海平面的海拔高度，或某个参考平面的相对高度，所以高程模型又叫地形模型。实际上地形模型不仅包含高程属性，还包含其它的地表形态属性，如坡度、坡向等。

数字地形模型是地形表面形态属性信息的数字表达，是带有空间位置特征和地形属性特征的数字描述。数字地形模型中地形属性为高程时称为数字高程模型（Digital Elevation Model，简称 DEM）。高程是地理空间中的第三维坐标。由于传统的地理信息系统的数据结构都是二维的，数字高程模型的建立是一个必要的补充。DEM 通常用地表规则网格单元构成的高程矩阵表示，广义的 DEM 还包括等高线、三角网等所有表达地面高程的数字表示。在地理信息系统中，DEM 是建立 DTM 的基础数据，其它的地形要素可由 DEM 直接或间接导出，称为“派生数据”，如坡度、坡向。

1．2 DEM的表示法

一个地区的地表高程的变化可以采用多种方法表达，用数学定义的表面或点、线、影像都可用来表示DEM，如图9-1所示。

1）数学方法

用数学方法来表达，可以采用整体拟合方法，即根据区域所有的高程点数据，用傅立叶级数和高次多项式拟合统一的地面高程曲面。也可用局部拟合方法，将地表复杂表面分成正方形规则区域或面积大致相等的不规则区域进行分块搜索，根据有限个点进行拟合形成高程曲面。

2）图形方法

（2．1）线模式

等高线是表示地形最常见的形式。其它的地形特征线也是表达地面高程的重要信息源，如山脊线、谷底线、海岸线及坡度变换线等。

（2．2）点模式

用离散采样数据点建立DEM是DEM建立常用的方法之一。数据采样可以按规则格网采样，可以是密度一致的或不一致的；可以是不规则采样，如不规则三角网、邻近网模型等；也可以有选择性地采样，采集山峰、洼坑、隘口、边界等重要特征点。

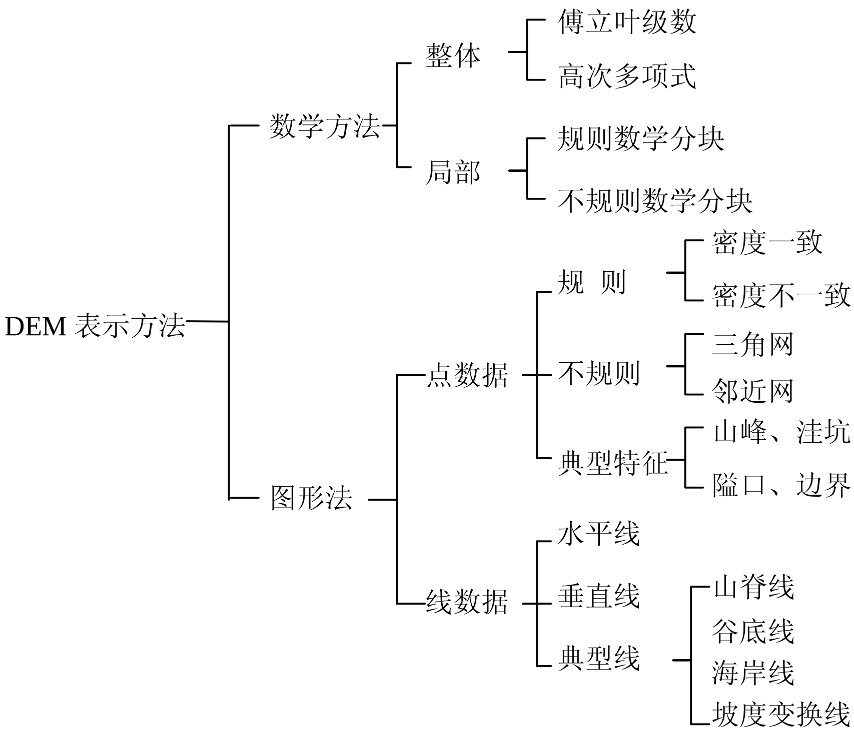


图9-1：DEM的表示方法

在地理信息系统中，DEM最主要的三种表示模型是：规则格网模型，等高线模型和不规则三角网模型。

2．DEM 的主要表示模型

2．1 规则格网模型

规则网格，通常是正方形，也可以是矩形、三角形等规则网格。规则网格将区域空间切分为规则的格网单元，每个格网单元对应一个数值。数学上可以表示为一个矩阵，在计算机实现中则是一个二维数组。每个格网单元或数组的一个元素，对应一个高程值，如图 9-2 所示。

91	78	63	50	53	63	44	55	43	25
94	81	64	51	57	62	50	60	50	35
100	84	66	55	64	66	54	65	57	42
103	84	66	56	72	71	58	74	65	47
96	82	66	63	80	78	60	84	72	49
91	79	66	66	80	80	62	86	77	56
86	78	68	69	74	75	70	93	82	57
80	75	73	72	68	75	86	100	81	56
74	67	69	74	62	66	83	88	73	53
70	56	62	74	57	58	71	74	63	45

图 9-2：格网 DEM

对于每个格网的数值有两种不同的解释。第一种是格网栅格观点，认为该格网单元的数值是其中所有点的高程值，即格网单元对应的地面面积内高程是均一的高度，这种数字高程模型是一个不连续的函数。第二种是点栅格观点，认为该网格单元的数值是网格中心点的高程或该网格单元的平均高程值，这样就需要用一种插值方法来计算每个点的高程。计算任何不是网格中心的数据点的高程值，使用周围 4 个中心点的高程值，采用距离加权平均方法进行计算，当然也可使用样条函数和克里金插值方法。

规则格网的高程矩阵，可以很容易地用计算机进行处理，特别是栅格数据结构的地理信息系统。它还可以很容易地计算等高线、坡度坡向、山坡阴影和自动提取流域地形，使得它成为 DEM 最广泛使用的格式，目前许多国家提供的 DEM 数据都是以规则格网的数据矩阵形式提供的。格网 DEM 的缺点是不能准确表示地形的结构和细部，为避免这些问题，可采用附加地形特征数据，如地形特征点、山脊线、谷底线、断裂线，以描述地形结构。

格网 DEM 的另一个缺点是数据量过大，给数据管理带来了不方便，通常要进行压缩存储。DEM 数据的无损压缩可以采用普通的栅格数据压缩方式，如游程编码、块码等，但是由于 DEM 数据反映了地形的连续起伏变化，通常比较“破碎”，普通压缩方式难以达到很好的效果；因此对于网格 DEM 数据，可以采用哈夫曼编码进行无损压缩；有时，在牺牲细节信息的前提下，可以对网格 DEM 进行有损压缩，通常的有损压缩大都是基于离散余弦变换（Discrete Cosine Transformation，DCT）或小波变换（Wavelet Transformation）的，由于小波变换具有较好的保持细节的特性，近年来将小波变换应用于 DEM 数据处理的研究较多。

2．2 等高线模型

等高线模型表示高程，高程值的集合是已知的，每一条等高线对应一个已知的高程值，这样一系列等高线集合和它们的高程值一起就构成了一种地面高程模型。如图 9-3 所示。

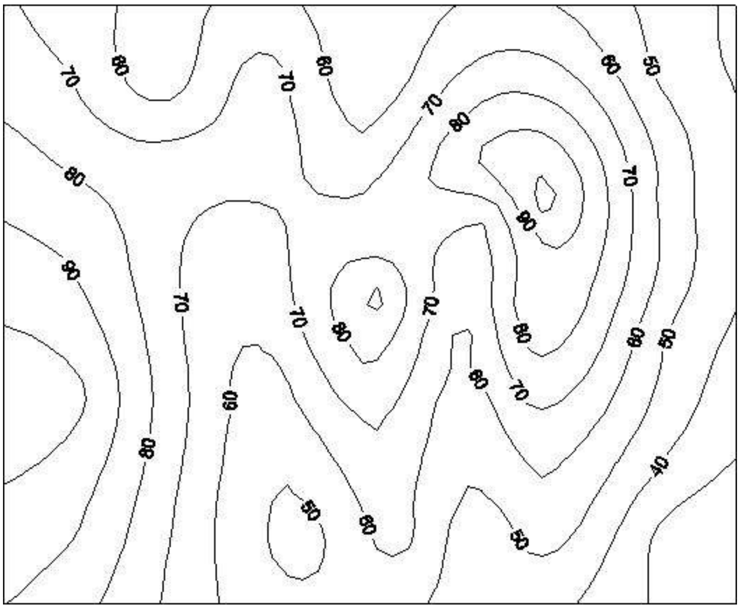


图 9-3：等高线

等高线通常被存成一个有序的坐标点对序列，可以认为是一条带有高程值属性的简单多边形或多边形弧段。由于等高线模型只表达了区域的部分高程值，往往需要一种插值方法来计算落在等高线外的其它点的高程，又因为这些点是落在两条等高线包围的区域内，所以，通常只使用外包的两条等高线的高程进行插值。

等高线通常可以用二维的链表来存储。另外的一种方法是用图来表示等高线的拓扑关系，将等高线之间的区域表示成图的节点，用边表示等高线本身。此方法满足等高线闭合或与边界闭合、等高线互不相交两条拓扑约束。这类图可以改造成一种无圈的自由树。下图为一个等高线图和它相应的自由树（图 9-4）。其它还有多种基于图论的表示方法。

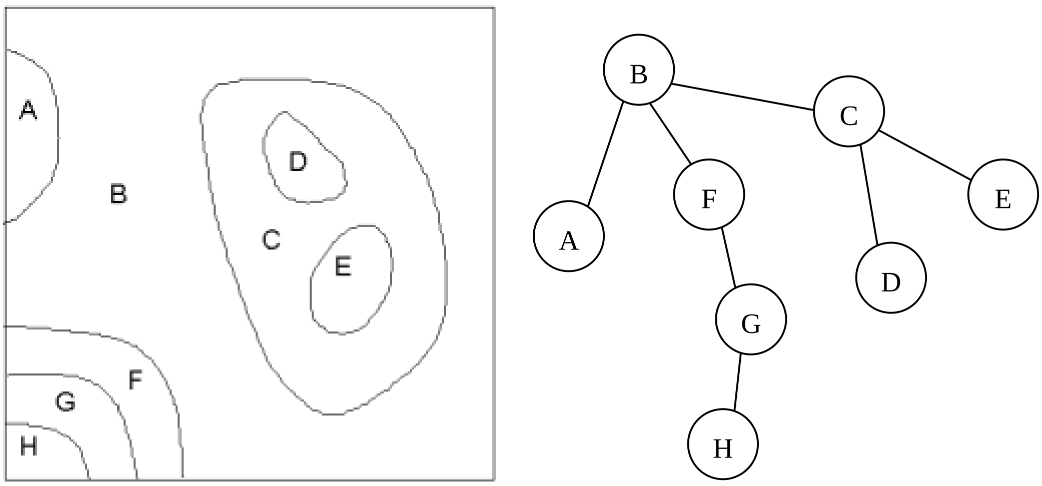


图 9-4：等高线和相应的自由树

2．3 不规则三角网（TIN）模型

尽管规则格网 DEM 在计算和应用方面有许多优点，但也存在许多难以克服的缺陷：

1）在地形平坦的地方，存在大量的数据冗余；

2）在不改变格网大小的情况下，难以表达复杂地形的突变现象；

3）在某些计算，如通视问题，过分强调网格的轴方向。

不规则三角网（Triangulated Irregular Network, TIN）是另外一种表示数字高程模型的方法[Peuker 等，1978]，它既减少规则格网方法带来的数据冗余，同时在计算（如坡度）效率方面又优于纯粹基于等高线的方法。

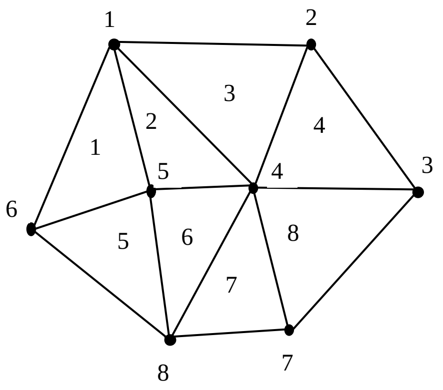
TIN 模型根据区域有限个点集将区域划分为相连的三角面网络，区域中任意点落在三角面的顶点、边上或三角形内。如果点不在顶点上，该点的高程值通常通过线性插值的方法得到（在边上用边的两个顶点的高程，在三角形内则用三个顶点的高程）。所以 TIN 是一个三维空间的分段线性模型，在整个区域内连续但不可微。

TIN 的数据存储方式比格网 DEM 复杂，它不仅要存储每个点的高程，还要存储其平面坐标、节点连接的拓扑关系，三角形及邻接三角形等关系。TIN 模型在概念上类似于多边形网络的矢量拓扑结构，只是 TIN 模型不需要定义“岛”和“洞”的拓扑关系。

有许多种表达 TIN 拓扑结构的存储方式，一个简单的记录方式是：对于每一个三角形、边和节点都对应一个记录，三角形的记录包括三个指向它三个边的记录的指针；边的记录有四个指针字段，包括两个指向相邻三角形记录的指针和它的两个顶点的记录的指针；也可以直接对每个三角形记录其顶点和相邻三角形（图 9-5）。每个节点包括三个坐标值的字段，分别存储 X，X，Z 坐标。这种拓扑网络结构的特点是对于给定一个三角形查询其三个顶点高程和相邻三角形所用的时间是定长的，在沿直线计算地形剖面线时具有较高的效率。当然可以在此结构的基础上增加其它变化，以提高某些特殊运算的效率，例如在顶点的记录里增加指向其关联的边的指针。

1	X	Y	Z
2	X	Y	Z
3	X	Y	Z
4	X	Y	Z
5	X	Y	Z
6	X	Y	Z
7	X	Y	Z
8	X	Y	Z

点文件



		顶点				邻接三角形		
1		1	5	6		2	5	X
2		1	4	5		1	3	6
3		1	2	4		X	4	2
4		2	3	4		3	X	8
5		5	6	8		1	X	6
6		4	5	8		2	5	7
7		4	7	8		6	8	X
8		3	4	7		4	7	X

三角形文件

图 9-5：三角网的一种存储方式

不规则三角网数字高程由连续的三角面组成，三角面的形状和大小取决于不规则分布的测点，或节点的位置和密度。不规则三角网与高程矩阵方法不同之处是随地形起伏变化的复杂性而改变采样点的密度和决定采样点的位置，因而它能够避免地形平坦时的数据冗余，又能按地形特征点如山脊、山谷线、地形变化线等表示数字高程特征。

2．4 层次模型

层次地形模型（Layer of Details，LOD）是一种表达多种不同精度水平的数字高程模型。大多数层次模型是基于不规则三角网模型的，通常不规则三角网的数据点越多精度越高，数据点越少精度越低，但数据点多则要求更多的计算资源。所以如果在精度满足要求的情况下，最好使用尽可能少的数据点。层次地形模型允许根据不同的任务要求选择不同精度的地形模型。层次模型的思想很理想，但在实际运用中必须注意几个重要的问题：

1）层次模型的存储问题，很显然，与直接存储不同，层次的数据必然导致数据冗余。

2）自动搜索的效率问题，例如搜索一个点可能先在最粗的层次上搜索，再在更细的层次上搜索，直到找到该点。

3）三角网形状的优化问题，例如可以使用 Delaunay 三角剖分。

4）模型可能允许根据地形的复杂程度采用不同详细层次的混合模型，例如，对于飞行模拟，近处时必须显示比远处更为详细的地形特征。

5）在表达地貌特征方面应该一致，例如，如果在某个层次的地形模型上有一个明显的山峰，在更细层次的地形模型上也应该有这个山峰。

这些问题目前还没有一个公认的最好的解决方案，仍需进一步深入研究。

3．DEM 模型之间的相互转换

在实际应用中，DEM 模型之间可以相互转换。大部分 DEM 数据都是规则格网 DEM，但由于规则格网 DEM 的数据量大而不便存储，也可能由于某些分析计算需要使用 TIN 模型的 DEM，如进行通视分析。此时需要将格网 DEM 转成 TIN 模型的 DEM。反之，如果已有 TIN 模型的 DEM 数据，为满足某种应用的需要，也需要转成规则格网的 DEM。

3．1 不规则点集生成 TIN

对于不规则分布的高程点，可以形式化地描述为平面的一个无序的点集 P，点集中每个点 p 对应于它的高程值。将该点集转成 TIN，最常用的方法是 Delaunay 三角剖分方法。生成 TIN 的关键是 Delaunay 三角网的产生算法，下面先对 Delaunay 三角网和它的偶图 Voronoi 图作简要的描述。

Voronoi 图，又叫泰森多边形或 Dirichlet 图，它由一组连续多边形组成，多边形的边界是由连接两邻点线段的垂直平分线组成。N 个在平面上有区别的点，按照最近邻原则划分平面：每个点与它的最近邻区域相关联。Delaunay 三角形是由与相邻 Voronoi 多边形共享一条边的相关点连接而成的三角形。Delaunay 三角形的外接圆圆心是与三角形相关的 Voronoi 多边形的一个顶点。Delaunay 三角形是 Voronoi 图的偶图，如图 9-6 所示。

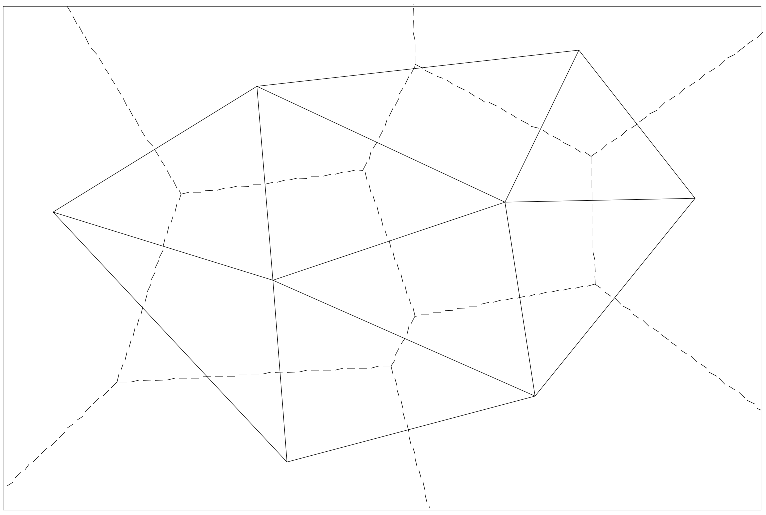


图 9-6：Delaunay 三角网与 Voronoi 图

对于给定的初始点集 P，有多种三角网剖分方式，而 Delaunay 三角网有以下特性：

1）其 Delaunay 三角网是唯一的；

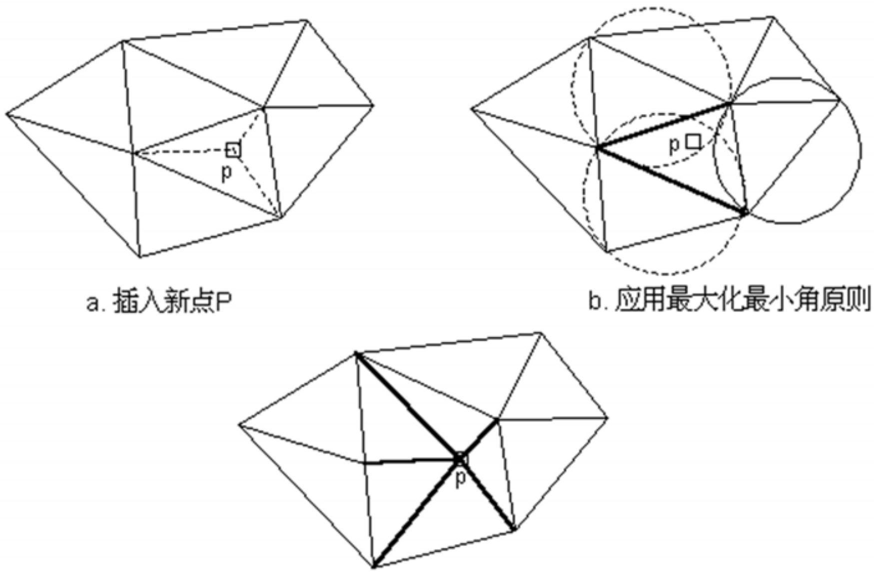
2）三角网的外边界构成了点集 P 的凸多边形“外壳”；

3）没有任何点在三角形的外接圆内部，反之，如果一个三角网满足此条件，那么它就是 Delaunay 三角网。

4）如果将三角网中的每个三角形的最小角进行升序排列，则 Delaunay 三角网的排列得到的数值最大，从这个意义上讲，Delaunay 三角网是“最接近于规则化”的三角网。

下面简要介绍 Delaunay 三角形产生的基本准则：

Delaunay 三角形产生准则的最简明的形式是：任何一个 Delaunay 三角形的外接圆的内部不能包含其它任何点[Delaunay 1934]。Lawson[1972]提出了最大化最小角原则：每两个相邻的三角形构成的凸四边形的对角线，在相互交换后，六个内角的最小角不再增大。Lawson [1977]又提出了一个局部优化过程 LOP（Local Optimization Procedure）方法。如图 9-7 所示。先求出包含新插入点 *p* 的外接圆的三角形，这种三角形称为影响三角形（Influence Triangulation）。删除影响三角形的公共边（图 b 中粗线），将 *p* 与全部影响三角形的顶点连接，完成 *p* 点在原 Delaunay 三角形中的插入。



a. 插入新点P　　b. 应用最大化最小角原则

c. 修改后的狄洛尼三角形

图 9-7：向 Delaunay 三角形中插入点

将该点集转成 TIN，最常用的方法是 Delaunay 三角剖分方法，生成过程分两步完成：

1）利用 P 中点集的平面坐标产生 Delaunay 三角网；

2）给 Delaunay 三角形中的节点赋予高程值。

3．2 格网 DEM 转成 TIN

格网 DEM 转成 TIN 可以看作是一种规则分布的采样点生成 TIN 的特例，其目的是尽量减少 TIN 的顶点数目，同时尽可能多地保留地形信息，如山峰、山脊、谷底和坡度突变处。规则格网 DEM 可以简单地生成一个精细的规则三角网，针对它有许多算法，绝大多数算法都有两个重要的特征：

1）筛选要保留或丢弃的格网点；

2）判断停止筛选的条件。

其中两个代表性的方法算法是保留重要点法和启发丢弃法。

3．2．1 保留重要点法

该方法是一种保留规则格网 DEM 中的重要点来构造 TIN 的方法[Chen、Gauvara（1987）]。它是通过比较计算格网点的重要性，保留重要的格网点。重要点（VIP，Very Important Point）是通过 3*3 的模板来确定的，根据八邻点的高程值决定模板中心是否为重要点。格网点的重要性是通过它的高程值与 8 邻点高程的内插值进行比较，当差分超过某个阈值的格网点保留下来。被保留的点作为三角网顶点生成 Delaunay 三角网。如图 9-8 所示，由 3*3 的模板得到中心点 P 和 8 邻点的高程值，计算中心点 P 到直线 AE，CG，BF，DH 的距离，图右图表示，再计算 4 个距离的平均值。如果平均值超过阈值，P 点为重要点，则保留，否则去除 P 点。

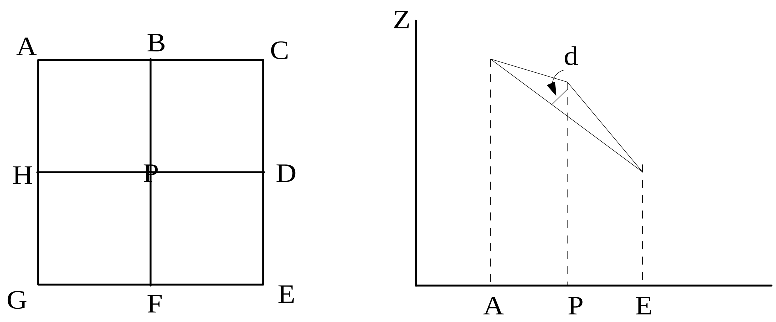


图 9-8：VIP 方法示意

3．2．2 启发丢弃法（DH—Drop Heuristic）

该方法将重要点的选择作为一个优化问题进行处理。算法是给定一个格网 DEM 和转换后 TIN 中节点的数量限制，寻求一个 TIN 与规则格网 DEM 的最佳拟合。首先输入整个格网 DEM，迭代进行计算，逐渐将那些不太重要的点删除，处理过程直到满足数量限制条件或满足一定精度为止。具体过程如下（图 9-9）：

1）算法的输入是 TIN，每次去掉一个节点进行迭代，得到节点越来越少的 TIN。很显

然，可以将格网 DEM 作为输入，此时所有格网点视为 TIN 的节点，其方法是将格网中 4 个节点的其中两个相对节点连接起来，这样将每个格网剖分成两个三角形。

2）取 TIN 的一个节点 O 及与其相邻的其它节点，如图 9-9 所示，O 的邻点（称 Delaunay 邻接点）为 A，B，C，D，使用 Delaunay 三角构造算法，将 O 的邻点进行 Delaunay 三角形重构，图 9-9 中实线所示。

3）判断该节点O位于哪个新生成的Delaunay三角形中，如图 9-9 为三角形BCE。计算O点的高程和过O点与三角形BCE交点O’的高程差d。若高程差d大于阈值d_e，则O点为重要点，保留，否则，可删除。d_e为阈值。

4）对 TIN 中所有的节点，重复进行上述判断过程。

5）直到TIN中所有的节点满足条件$d>d_e$，结束。

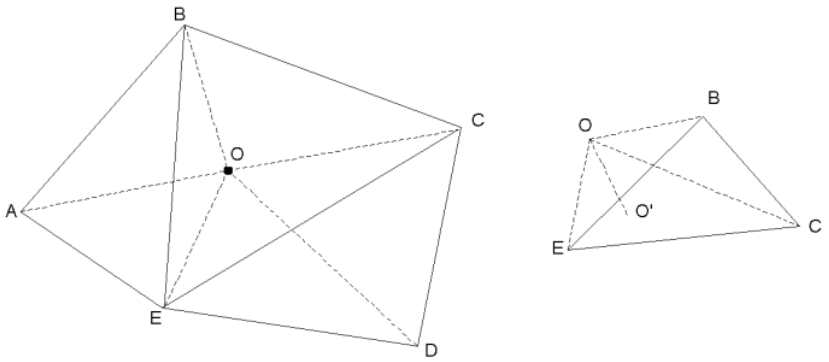


图 9-9：DH 方法转换格网 DEM 成 TIN

（左图虚线为以 O 为中心的 Delaunay 三角形，实线为新生成的 Delaunay 三角形；

右图为高差的计算[注意：此图描述了三维空间]）

两种方法相比较[Lee，1991]，VIP 方法在保留关键网格点方面（顶点、凹点）最好；DH 方法在每次丢弃数据点时确保信息丢失最少，但要求计算量大。各种方法各有利弊，实际应用中根据不同的需要，如检测极值点，高效存储，最小误差，可以选择使用不同的方法。

3．3 等高线转成格网 DEM

表示地形的最常见的线模式是一系列描述高程曲线的等高线。由于现有地图大多数都绘有等高线，这些地图便是数字高程模型的现成数据源，可以将纸面等高线图扫描后，自动获取 DEM 数据。由于数字化的等高线不适合于计算坡度或制作地貌渲染图等地形分析，因此，必须要把数字化等高线转为格网高程矩阵。

使用局部插值算法，如距离倒数加权平均或克里金插值算法*，可以将数字化等高线数据转为规则格网的DEM数据，但插值的结果往往会出现一些许多不令人满意的结果，而且数字化等高线时越小心，采样点越多，问题越严重。问题不在于计算插值权重系数的理论假设，也不在于平滑等高线是真实地形的反映的假设，而在于估计未知格网点的高程要在一个半径范围内搜索落在其中的已知点数据，再计算它的加权平均值。如果搜索到的点都具有相同的高程，那待插值点的高程也同为此高程值。结果导致在每条等高线周围的狭长区域内具有与等高线相同的高程，出现了“阶梯”地形。当低海拔平原地区等高线距离更远时，搜索到一条等高线上的数据的可能性就越大，问题更严重。以带“阶梯”地形的DEM为基础，计算坡度往往会出现不自然的条斑状分布模式（图 9-10）。

* 见“空间分析”一章中的“空间插值”节。

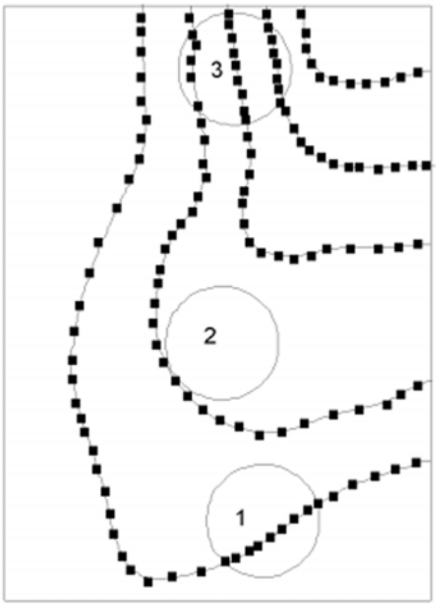
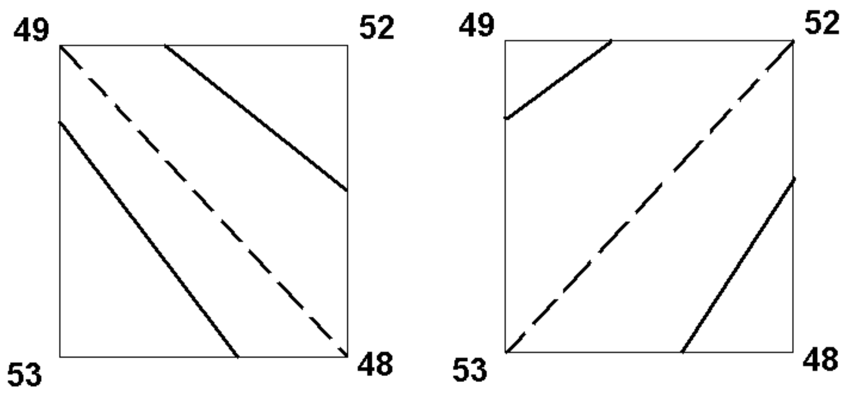


图 9-10：等值线插值造成“阶梯地形”的原因

最好的解决方法是使用针对等高线插值的专用方法。如果没有合适的方法，最好把等高线数据点减少到最少，增加标识山峰、山脊、谷底和坡度突变的数据点，同时使用一个较大的搜索窗口。

3．4 利用格网 DEM 提取等高线

在利用格网 DEM 生成等高线时，需要将其中的每个点视为一个几何点，而不是一个矩形区域，这样可以根据格网 DEM 中相邻四个点组成四边形进行等高线跟踪。其方法类似于后面描述的利用 TIN 提取等高线。实际上，也可以将每个矩形分割成为两个三角形，并应用 TIN 提取等高线算法，但是由于矩形有两种划分三角形的方法，在某些情况下，会生成不同的等高线（图 9-11），这时需要根据周围的情况进行判断并决定取舍。



（a） (b)

图 9-11：由于三角形划分不同造成生成等高线的不同

在格网 DEM 提取等高线中，除了划分为三角形之外，也可以直接使用四边形跟踪等高线。但是在图 9-11 所示的情形中，仍会出现等高线跟踪的二义性，即对于每个四边形，有两条等高线的离去边。进行取舍判断的方法一般是计算距离，距离近的连线方式优于距离远的连线方式。在图 9-11 种，就要采用（b）图所示的跟踪方式。

格网 DEM 提取等高线另一个值得注意的问题是，如果一些网格点的数值恰好等于要提取的等高线的数值，会使判断过程变得复杂，并且会生成不闭合的等高线，一般的解决办法是将这些网格点的数值增加一个小的偏移量。

3．5 TIN 转成格网 DEM

TIN转成格网DEM可以看作普通的不规则点生成格网DEM的过程。方法是按要求的分辨率大小和方向生成规则格网，对每一个格网搜索最近的TIN数据点，按线性或非线性插值函数计算格网点高程*。

4．DEM 的建立

为了建立 DEM，必需量测一些点的三维坐标，这就是 DEM 数据采集。

4．1 DEM 数据采集方法

1）地面测量

利用自动记录的测距经纬仪（常用电子速测经纬仪或全站经纬仪）在野外实测。这种速测经纬仪一般都有微处理器，可以自动记录和显示有关数据，还能进行多种测站上的计算工作。其记录的数据可以通过串行通讯，输入计算机中进行处理。

2）现有地图数字化

利用数字化仪对已有地图上的信息（如等高线）进行数字化的方法，目前常用的数字化仪有手扶跟踪数字化仪和扫描数字化仪。

3）空间传感器

利用全球定位系统 GPS，结合雷达和激光测高仪等进行数据采集。

4）数字摄影测量方法

这是 DEM 数据采集最常用的方法之一。利用附有的自动记录装置（接口）的立体测图仪或立体坐标仪、解析测图仪及数字摄影测量系统，进行人工、半自动或全自动的量测来获取数据。

4．2 数字摄影测量获取 DEM

数字摄影测量方法是空间数据采集最有效的手段，它具有效率高、劳动强度低的优点。数据采样可以全部由人工操作，通常费时且易于出错；半自动采样可以辅助操作人员进行采样，以加快速度和改善精度，通常是由人工控制高程 Z，由机器自动控制平面坐标 X,Y 的驱动；全自动方法利用计算机视觉代替人眼的立体观测，速度虽然快，但精度较差。

人工或半自动方式的数据采集，数据的记录可分为“点模式”或“流模式”，前者根据控制信号记录静态量测数据，后者是按一定规律连续地记录动态的量测数据。

摄影测量方法用于生产 DEM，数据点的采样方法根据产品的要求不同而异。沿等高线、断面线、地性线进行采样往往是有目的的采样。而许多产品要求高程矩阵形式，所以基于规

* 具体的计算方法见第五节第二部分。

则格网或不规则格网点的面采样是必须的，这种方式与其它空间属性的采样方式一样，只是采样密度高一些。

1）沿等高线采样

在地形复杂及陡峭地区，可采用沿等高线跟踪方式进行数据采集，而在平坦地区，则不宜采用沿等高线采样。沿等高线采样时可按等距离间隔记录数据或按等时间间隔记录数据方式进行。采用后一种方式，由于在等高线曲率大的地方跟踪速度较慢，因而采集的点较密集，而在等高线较平直的地方跟踪速度快，采集的点较稀疏，故只要选择恰当的时间间隔，所记录的数据就能很好地描述地形，又不会有太多的数据。

2）规则格网采样

利用解析测图仪在立体模型中按规则矩形格网进行采样，直接构成规则格网 DEM。当系统驱动测标到格网点时，会按预先选定的参数停留一短暂时间（如 0.2 秒），供作业人员精确测量。该方法的优点是方法简单、精度高、作业效率也较高；缺点是对地表变化的尺度的灵活性较差，可能会丢失特征点。

3）渐进采样（Progressive Sampling）

渐进采样方法的目的是使采样点分布合理，即平坦地区样点少，地形复杂区的样点较多。渐进采样首先按预定比较稀疏的间隔进行采样，获得一个较稀疏的格网，然后分析是否需要对格网进行加密，如图 9-12 所示。判断加密的方法可利用高程的二阶差分是否超过了给定的阈值；或利用相邻的三点拟合一条二次曲线，计算两点间中点的二次内插值与线性内插值之差，判断是否超过阈值。当超过阈值时，则对格网加密采样，然后对较密的格网进行同样的判断处理，直至不再超限或达到预先给定的加密次数（或最小格网间隔），然后再对其它格网进行同样的处理。

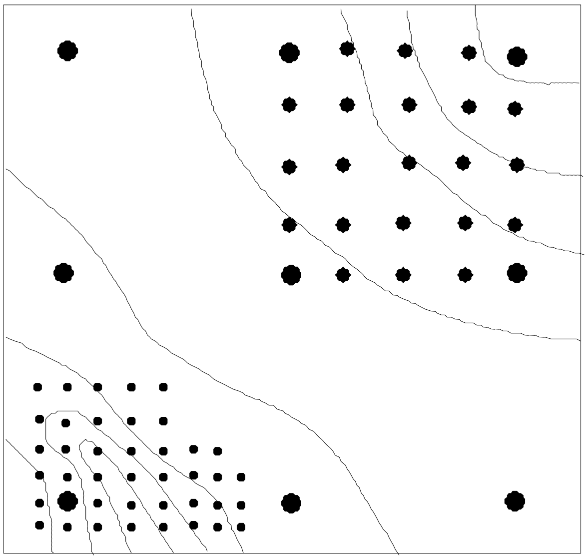


图 9-12：渐进采样

4）选择采样

为了准确地反映地形，可根据地形特征进行选择采样，例如沿山脊线、山谷线、断裂线进行采集以及离散碎部点（如山顶）的采集。这种方法获取的数据尤其适合于不规则三角网 DEM 的建立。

5）混合采样

为了同步考虑采样的效率与合理性，可将规则采样（包括渐进采样）与选择性采样结合进行混合采样，即在规则采样的基础上再进行沿特征线、点采样。为了区别一般的数据点和特征点，应当给不同的点以不同的特征码，以便处理时可按不同的方式进行。利用混合采样

可建立附加地形特征的规则格网 DEM，也可建立附加特征的不规则三角网 DEM。

6）自动化 DEM 数据采集

上述方法均是基于解析测图仪或机助制图系统利用半自动的方法进行 DEM 数据采集，现在已经可以利用自动化测图系统进行完全自动化的 DEM 数据采集。此时可按像片上的规则格网利用数字影像匹配进行数据采集。

最后数字摄影测量获取的 DEM 数据点都要按一定插值方法转成规则格网 DEM 或规则三角网 DEM 格式数据。

4．3 DEM 数据质量控制

数据采集是 DEM 的关键问题，研究结果表明，任何一种 DEM 内插方法，均不能弥补取样不当所造成的信息损失。数据点太稀会降低 DEM 的精度；数据点过密，又会增大数据量、处理的工作量和不必要的存储量。这需要在 DEM 数据采集之前，按照所需的精度要求确定合理的取样密度，或者在 DEM 数据采集过程中根据地形复杂程度动态调整采样点密度。

由于很多 DEM 数据来源于地形图，所以 DEM 的精度决不会高于原始的地形图。例如 U.S.G.S.用数字化的等高线图，通过线性插值生产的最精确的 DEM 的最大均方误差（RMSE）为等高线间距的一半，最大误差不大于两个等高线间距。通常用某种数学拟合曲面生产的 DEM，往往存在未知的精度问题，即使是正式出版的地形图同样存在某种误差，所以在生产和使用 DEM 时应该注意到它的误差类型。

DEM 的数据质量可以参考美国 U.S.G.S.的分级标准，共分为三级：第一级，最大绝对垂直误差 50 米、最大相对垂直误差 21 米，绝大多数 7.5 分幅产品属于第一级；第二级 DEM 数据对误差进行了平滑和修改处理，数字化等高线插值生产的 DEM 属于第二级，最大误差为两个等间距，最大均方误差为半个等间距；第三级 DEM 数据最大误差为一个等间距，最大均方误差为三分之一个等间距。

5．DEM 的分析和应用

5．1 格网 DEM 应用

5．1．1 地形曲面拟合

DEM 最基础的应用是求 DEM 范围内任意点的高程，在此基础上进行地形属性分析。由于已知有限个格网点的高程，可以利用这些格网点高程拟合一个地形曲面，推求区域内任意点的高程。曲面拟合方法可以看作是一个已知规则格网点数据进行空间插值的特例，距离倒数加权平均方法，克里金插值方法，样条函数等插值方法均可采用。

5．1．2 立体透视图

从数字高程模型绘制透视立体图是 DEM 的一个极其重要的应用。透视立体图能更好地反映地形的立体形态，非常直观。与采用等高线表示地形形态相比有其自身独特的优点，更接

近人们的直观视觉。特别是随着计算机图形处理工作的增强以及屏幕显示系统的发展，使立体图形的制作具有更大的灵活性，人们可以根据不同的需要，对于同一个地形形态作各种不同的立体显示。例如局部放大，改变高程值 Z 的放大倍率以夸大立体形态；改变视点的位置以便从不同的角度进行观察，甚至可以使立体图形转动，使人们更好地研究地形的空间形态。

从一个空间三维的立体的数字高程模型到一个平面的二维透视图，其本质就是一个透视变换。将“视点”看作为“摄影中心”，可以直接应用共线方程从物点（*X*，*Y*，*Z*）计算“像点”坐标（*X*, *Y*)。透视图中的另一个问题是“消隐”的问题，即处理前景挡后景的问题。

调整视点、视角等各个参数值，就可从不同方位、不同距离绘制形态各不相同的透视图制作动画。计算机速度充分高时，就可实时地产生动画 DTM 透视图。

5．1．3 通视分析

通视分析有着广泛的应用背景。典型的例子是观察哨所的设定，显然观察哨的位置应该设在能监视某一感兴趣的区域，视线不能被地形挡住。这就是通视分析中典型的点对区域的通视问题。与此类似的问题还有森林中火灾监测点的设定，无线发射塔的设定等。有时还可能对不可见区域进行分析，如低空侦察飞机在飞行时，要尽可能躲避敌方雷达的捕捉，飞行显然要选择雷达盲区飞行。通视问题可以分为五类[Lee,J.（1991）]：

1）已知一个或一组观察点，找出某一地形的可见区域。

2）欲观察到某一区域的全部地形表面，计算最少观察点数量。

3）在观察点数量一定的前提下，计算能获得的最大观察区域。

4）以最小代价建造观察塔，要求全部区域可见。

5）在给定建造代价的前提下，求最大可见区。

根据问题输出维数的不同，通视可分为点的通视，线的通视和面的通视。点的通视是指计算视点与待判定点之间的可见性问题；线的通视是指已知视点，计算视点的视野问题；区域的通视是指已知视点，计算视点能可视的地形表面区域集合的问题。基于格网 DEM 模型与基于 TIN 模型的 DEM 计算通视的方法差异很大。

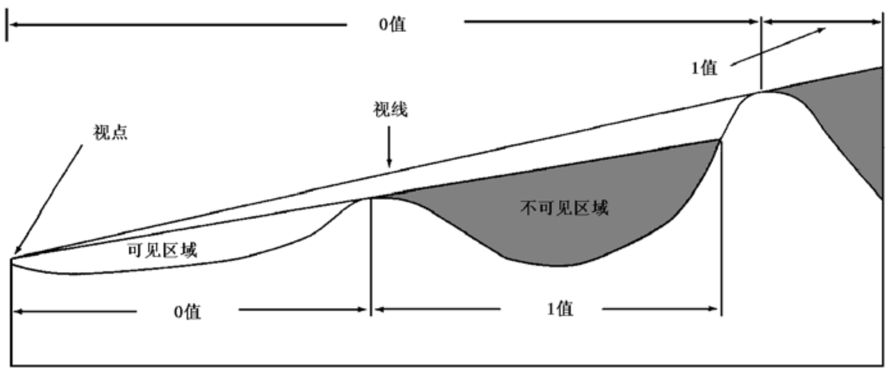


图 9-13：通视分析，图上灰色区域为不可见区域

1）点对点通视

基于格网 DEM 的通视问题，为了简化问题，可以将格网点作为计算单位。这样点对点的通视问题简化为离散空间直线与某一地形剖面线的相交问题。(图 9-13）

已知视点V的坐标为（x_0,y_0,z_0），以及P点的坐标（x_1,y_1,z_1）。DEM为二维数组Z[M][N]，则V为（$m_0,n_0,Z[m_0,n_0]$），P为（$m_1,n_1,Z[m_1,n_1]$）。计算过程如下：

（1．1）使用 Bresenham 直线算法，生成 V 到 P 的投影直线点集{x , y}，K=||{x , y}||, 并得到直线点集{x , y}对应的高程数据{Z[k], (k=1,...K-1)}，这样形成 V 到 P 的 DEM 剖面曲

线。

（1．2）以 V 到 P 的投影直线为 X 轴，V 的投影点为原点，求出视线在 X-Z 坐标系的直线方程：

$$H[k]=\frac{Z[m_0][n_0]-Z[m_1][n_1]}{K}\cdot k+Z[m_0][n_0] \quad （0<k<K）$$

K 为 V 到 P 投影直线上离散点数量。

（1．3）比较数组 H[k]与数组 Z[k]中对应元素的值，如果

$$\forall k, k\in[1, K-1]$$

存在 Z[k]>H[k]，则 V 与 P 不可见，否则可见。

2）点对线通视

点对线的通视，实际上就是求点的视野。应该注意的是，对于视野线之外的任何一个地形表面上的点都是不可见的，但在视野线内的点有可能可见，也可能不可见。基于格网 DEM 点对线的通视算法如下：

（2．1）设 P 点为一沿着 DEM 数据边缘顺时针移动的点，与计算点对点的通视相仿，求出视点到 P 点投影直线上点集{x, y}，并求出相应的地形剖面{x, y, Z(x, y)}。

（2．2）计算视点至每个 $p_k\in\{x,y,z(x,y)\}, k=1,2,\ldots K-1$ 与 Z 轴的夹角 β_k：

$$\beta_k=arctg\left(\frac{k}{Z_{pk}-Z_{vp}}\right)$$

（2．3）求得 $\alpha=\min\{\beta_k\}$。α 对应的点就为视点视野线的一个点。

（2．4）移动 P 点，重复以上过程，直至 P 点回到初始位置，算法结束。

3）点对区域通视

点对区域的通视算法是点对点算法的扩展。与点到线通视问题相同，P 点沿数据边缘顺时针移动。逐点检查视点至 P 点的直线上的点是否通视。一个改进的算法思想是，视点到 P 点的视线遮挡点，最有可能是地形剖面线上高程最大的点。因此，可以将剖面线上的点按高程值进行排序，按降序依次检查排序后每个点是否通视，只要有一个点不满足通视条件，其余点不再检查。点对区域的通视实质仍是点对点的通视，只是增加了排序过程。

5．1．4 流域特征地貌提取与地形自动分割

地形因素是影响流域地貌、水文、生物等过程的重要因子，地形属性的空间分布特征一直是人们用于描述这些空间过程变化的重要指标。高精度 DEM 数据和高分辨率、高光谱、多周期的遥感影像，为人们定量描述流域空间变化过程提供了日益丰富的数据源，而且人们对流域地貌、水文和生物等过程空间变化机理理解的不断加深，可以说人类已经进入了一个“空间模拟”的时代。基于 DEM 数据自动提取流域地貌特征和进行流域地形自动分割是进行流域空间模拟的基础技术。

基于格网 DEM 自动提取流域特征地貌和进行地形自动分割技术主要包括两个方面：1）流域地貌形态结构定义，定义能反映流域结构的特征地貌，建立格网 DEM 对应的微地貌特征。2）特征地貌自动提取和地形自动分割算法。格网 DEM 数据是一些离散的高程点数据，每个数据本身不能反映实际地表的复杂性。为了从格网 DEM 数据中得到流域地貌形态结构，

必须采用一个清晰的流域地貌结构模型，然后针对该结构模型设计自动提取算法。

1）流域结构定义

可以使用一个具有根的树状图来描述流域结构[Shreve]，目前绝大多数算法都沿用这一描述方法。在此结构中主要包括三个部分，即结点集、界线集和汇流区集。如图 9-14 所示。

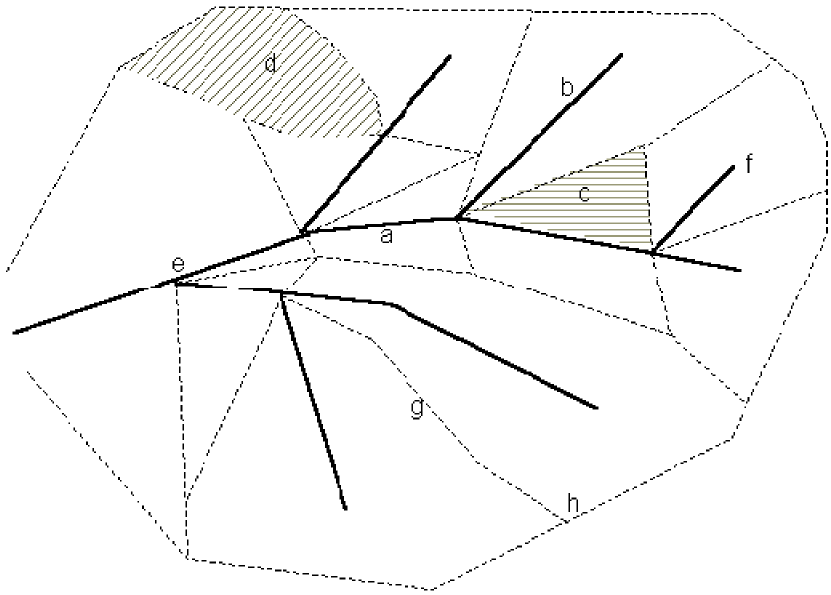


图 9-14：流域结构

（a.内部沟谷段 b. 外部沟谷段 c. 内部汇流区 d. 外部汇流区 e. 沟谷结点 f. 汇流源点 g. 分水线段 h. 分水线源点）

其具体内容包括几个概念：

1）沟谷线段：一条具有两侧汇流区的线段；

2）分水线段：一条具有两侧分水区的线段；

3）沟谷结点：两条或两条以上沟谷线的交点；

4）分水线结点：两条或两条以上分水线的交点；

5）沟谷源点：沟谷的上游起点；

6）分水线源点：分水线与流域边界的交点；

7）内部汇流区：汇流区边界不包含流域部分边界的汇流区；

8）外部汇流区：汇流区边界包括部分流域边界的汇流区。

沟谷结点和沟谷源点共同组成沟谷结点集，所有的沟谷段组成沟谷段集，形成沟谷网络；所有的分水线组成分水线段集，形成分水线网络。沟谷段集和分水线段集共同把流域分割成一个汇流区集。

沟谷段是最小的沟谷单位，沟谷段可以分为内部沟谷段和外部沟谷段。内部沟谷段连接两个沟谷结点，外部沟谷段连接一个沟谷结点和沟谷源点。同样，分水线段是最小的分水线单位，也分为内部分水线段和外部分水线段。内部分水线段连接两个分水线结点，外部分水线段连接一个分水线结点和一个分水线源点。

汇流网络中每一沟谷段都有一个汇流区域，这些区域由分水线集控制。外部沟谷段有一个外部汇流区，内部沟谷段有两个内部汇流区，分布在内部沟谷段两侧。整个流域被分割成一个个子流域，每个子流域如同树状图上的一片“叶子”。

2）流域特征地貌自动提取和地形自动分割

特征地貌定义与提取：根据网格点高程与周围高程值的关系，将格网点分为坡地、洼地、分水线、谷地、阶地和鞍部等几类。先计算中心点与八邻点的高程差，然后对高程差进行排

序，再根据高程差序列的特性给中心点格网赋一个特征编码。然后通过一系列特征码的组合特征，用模式识别的方法，将格网点划分到已知的特征地貌类别。

山脊线和山谷线提取：山脊线和山谷线的自动探测实际上是凹点和凸点的自动搜索。较为简单的算子是 2*2 的局部算子。将算子在 DEM 数据中滑动，比较每个格网点与行和列上相邻格网点的高程，标出其中高程最小（探测山谷线）或高程最大（探测山脊线）的格网点。对整个 DEM 数据计算一遍后，剩下的未标记格网点就是山脊线或山谷线上的格网点。

流域地形自动分割：流域地形自动分割的目标是将整个流域分割成一个个子汇流区。大多数算法是利用 3*3 窗口计算流向和基于“溢流跟踪”算法确定汇流网络。算法过程如下：

（2．1）格网点流向定义

采用 3×3 窗口按 8 方向搜索计算最大坡向为各网格点的流向。分别为 8 方向赋不同的代码，如右图所示。每个格网有一个从 1 到 9 的数值，代表它流向相邻象元的方向，如该象元为凹点，则其值为 5（图 9-15）。

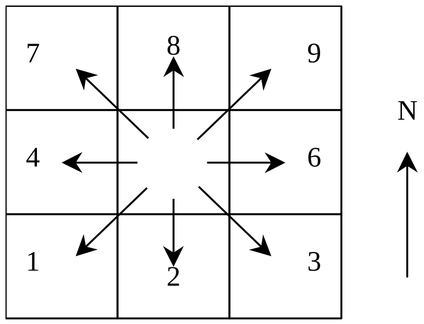


图 9-15：格网水流方向定义

几种例外情况的处理：

A．如果一个网格点的最大坡向格网点与之具有相同的高程值，且之前没有其它格网点流向这个相邻格网，则强制流向它。如果还有另外的格网点流向这个相邻格网，则当前格网点为凹点。

B．当两个或多个相邻格网点的最大坡向相等时，先比较各自相邻格网点坡向，如果仍没解决，继续比较相对格网点的坡向，决定赋一个流向。

C．对于具有相同高程值的区域则扩大搜索窗口半径，用 7×7 窗口，如果需要还可以使用更大窗口。

D．在 DEM 数据的外围加一圈高程值为 0 的格网点，强制其最大坡向流向研究区之外。当所有的格网点处理完毕后，生成一个编码 1—9 的流向图。

（2．2）凹点处理算法

由于凹点的存在，有一些流路不会流向流域出口，而是终止于凹点，所以在进行流域自动分割之前，还要对凹点进行处理。流域中凹点既可能是真实的凹点，也可能是由于插值误差造成的，所以不能使用简单的滤波或平滑函数，将凹点全部去除，目的是将凹点造成的断路连接到主沟谷网络。

搜索所有凹点的相邻最低点（有时可能有多个高程相等的最低点），作为凹点的溢出点，以溢出点为起点继续搜索比它的高程低或相等的邻点（已经搜索的点忽略），判断是否有比原凹点更低的格网点，如果没有则以该凹点的溢出点为起点，重复上述搜索过程；如果搜索到比原凹点低的格网点，将凹点和最低邻点的方向倒转。如图 9-15 所示：高程为 48 的点为一个凹点，搜索到高程最低的邻点为 49，以它为起点继续搜索，找到高程点 49，仍比原凹点高程高，则继续搜索，又找到另一个高程点 49，再找到高程 47 的点，比原凹点高程低，结束搜索，按搜索方向修改流向，如图 9-16 中实线箭头方向所示。

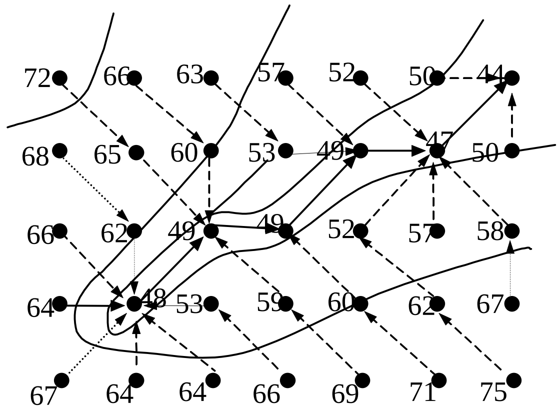


图 9-16：凹点处理

（2．3）提取汇流网络

根据修改后的流向图，给定一个点，所有流向它的格网点的总和就是该点的汇流区。计算方法是给定一个点，搜索 8 邻点，记录所有流向它的格网点的位置，然后再以找到的格网点为基点继续搜索记录流向它的格网点，直到没有新的汇流点为止，所有记录的格网点构成该点的汇流区。

通常沟谷的汇流区面积大于其它格网点的汇流区面积，可以通过设定一个阈值，将汇流区面积大于此阈值的格网点，标识为沟谷点。很明显，不同的阈值得到的沟谷网络的复杂性是不同的，这种方法虽然为确定沟谷网络的复杂性提供了灵活性，但也使得沟谷网络的确定具有太大的随意性。

得到沟谷网络后，可以对沟谷网络进行编码。首先对沟谷结点编码。从流域出口开始搜索遍历整个汇流网络，对每个沟谷段的上下游结点进行编码标识，标识值是沟谷段的编码值，并记录下这些结点的位置。其次，把沟谷段中的每个格网点标识为沟谷段的编码值。第三，根据沟谷段上游结点的类型判定沟谷段是内部沟谷段还是外部沟谷段。

（2．4）提取分水网络

递归搜索沟谷段中的每个格网点的汇流区，将汇流区的格网点赋为该沟谷段的标识值，形成各沟谷段的子汇流区。然后进行边界跟踪，提取子汇流区的边界线为分水线，得到分水线网络。最后，对沟谷网络和分水线网络及子汇流区进行拓扑编码，以完成流域地形的自动分割。

5．1．5 DEM 计算地形属性

由 DEM 派生的地形属性数据可以分为单要素属性和复合属性二种。前者可由高程数据直接计算得到，如坡度因子，坡向。后者是由几个单要素属性按一定关系组合成的复合指标，用于描述某种过程的空间变化，这种组合关系通常是经验关系，也可以使用简化的自然过程机理模型。

单要素地形属性通常可以很容易地使用计算机程序计算得到，包括：

1）坡度、坡向

坡度定义为水平面与局部地表之间的正切值。它包含两个成分：斜度——高度变化的最大值比率（常称为坡度）；坡向——变化比率最大值的方向。地貌分析还可能用到二阶差分凹率和凸率。比较通用的度量方法是：斜度用百分比度量，坡向按从正北方向起算的角度测量，凸度按单位距离内斜度的度数测量。

坡度和坡向的计算通常使用 3*3 窗口，窗口在 DEM 高程矩阵中连续移动后，完成整幅

图的计算。坡度的计算如下：

$$\tan\beta = \left[(\sigma_z/\sigma_x)^2 + (\sigma_z/\sigma_y)^2\right]^{1/2}$$

坡向计算如下：

$$\tan A = (-\sigma_z/\sigma_y)/(\sigma_z/\sigma_x) \qquad (-\pi < A < \pi)$$

为了提高计算速度和精度，GIS 通常使用二阶差分计算坡度和坡向，最简单的有限二阶差分法是按下式计算点 i,j 在 x 方向上的斜度：

$$(\sigma_z/\sigma_x)_{ij} = (z_{i+1,j} - z_{i-1,j})/2\sigma_x$$

式中 σ_x 是格网间距（沿对角线时 σ_x 应乘以 $\sqrt{2}$ ）。这种方法计算八各方向的斜度，运算速度也快得多。但地面高程得局部误差将引起严重得坡度计算误差，可以用数字分析方法来得到更好得结果，用数字分析方法计算东西方向得坡度公式如下：

$$(\sigma_z/\sigma_x)_{ij} = \left[(z_{i+1,j+1} + 2z_{i+1,j} + z_{i+1,j-1}) - (z_{i_1,j+1} + 2z_{i-1,j} + z_{i-1,j-1})\right]/8\sigma_x$$

同理可以写出其它方向的坡度计算公式。

2）面积、体积

（2．1）剖面积

根据工程设计的线路，可计算其与DEM各格网边交点P_i（X_i，Y_i，Z_i），则线路剖面积为

$$S = \sum_{i=1}^{n-1} \frac{Z_i + Z_{i+1}}{2} \cdot D_{i,i+1}$$

其中n为交点数；$D_{i,i+1}$为P_i与P_{i+1}之距离。同理可计算任意横断面及其面积。

（2．2）体积

DEM 体积由四棱柱（无特征的格网）与三棱柱体积进行累加得到，四棱柱体上表面用抛物双曲面拟合，三棱柱体上表面用斜平面拟合，下表面均为水平面或参考平面，计算公式分别为

$$\left.\begin{aligned} V_3 &= \frac{Z_1 + Z_2 + Z_3}{3} \cdot S_3 \\ V_4 &= \frac{Z_1 + Z_2 + Z_3 + Z_4}{4} \cdot S_4 \end{aligned}\right\}$$

其中S_3与S_4分别是三棱柱与四棱柱的底面积。

根据两个 DEM 可计算工程中的挖方、填方及土壤流失量。

3）表面积

对于含有特征的格网，将其分解成三角形，对于无特征的格网，可由 4 个角点的高程取平均即中心点高程，然后将格网分成 4 个三角形。由每一三角形的三个角点坐标（x_i，y_i，z_i）计算出通过该三个顶点的斜面内三角形的面积，最后累加就得到了实地的表面积。

5．2 三角网 DEM 分析应用

5．2．1 三角网内插

在建立 TIN 后，可以由 TIN 解求该区域内任意一点的高程。TIN 的内插与矩形格网的内插有不同的特点，其用于内插的点的检索比网格的检索要复杂。一般情况下仅用线性内插，即三角形三点确定的斜平面作为地表面，因而仅能保证地面连续而不能保证光滑。进行三角网内插，一般要经过以下几个步骤：

1）格网点的检索

给定一点的平面坐标P（x，y），要基于TIN内插该点的高程Z，首先要确定点P落在TIN的哪个三角形中。一般的做法是通过计算距离，得到据P点最近的点，设为Q_1。然后就要确定P所在的三角形。依次取出Q_1为顶点的三角形，判断P是否位于该三角形内。可利用P是否与该三角形每一顶点均在该顶点所对边的同侧（点的坐标分别代人该边直线方程所得的值符号相同）加以判断。若P不在以Q_1为顶点的任意一个三角形中，则取离P次最近的格网点，重复上述处理，直至取出P所在的三角形，即检索到用于内插P点高程的三个格网点。

2）高程内插

若P（x，y）所在的三角形为Δ$Q_1Q_2Q_3$，三顶点坐标为（x_1，y_1，z_1），（x_2，y_2，z_2）与（x_3，y_3，z_3），则由Q_1，Q_2与Q_3确定的平面方程为

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

或

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0$$

令

$$x_{21} = x_2 - x_1; \qquad x_{31} = x_3 - x_1$$
$$y_{21} = y_2 - y_1; \qquad y_{31} = y_3 - y_1$$
$$z_{21} = z_2 - z_1; \qquad z_{31} = z_3 - z_1;$$

则 P 点高程为

$$Z = Z_1 - \frac{(x - x_1)(y_{21} z_{31} - y_{31} z_{21}) + (y - y_1)(z_{21} x_{31} - z_{31} x_{21})}{x_{21} y_{31} - x_{31} y_{21}}$$

5．2．2 等高线追踪

基于 TIN 绘制等高线直接利用原始观测数据，避免了 DTM 内插的精度损失，因而等高线精度较高；对高程注记点附近的较短封闭等高线也能绘制；绘制的等高线分布在采样区域内

而并不要求采样区域有规则四边形边界。而同一高程的等高线只穿过一个三角形最多一次，因而程序设计也较简单。但是，由于 TIN 的存贮结构不同，等高线的具体跟踪算法跟踪也有所不同。

基于三角形搜索的等高线绘制算法如下：

对于记录了三角形表的 TIN，按记录的三角形顺序搜索。其基本过程如下：

1）对给定的等高线高程h，与所有网点高程z_i（i=1,2,…，n），进行比较，若$z_i=h$，则将z_i加上（或减）一个微小正数 $\varepsilon>0$（如 $\varepsilon=10^{-4}$），以使程序设计简单而又不影响等高线的精度。

2）设立三角形标志数组，其初始值为零，每一元素与一个三角形对应，凡处理过的三角形将标志置为 1，以后不再处理，直至等高线高程改变。

3）按顺序判断每一个三角形的三边中的两条边是否有等高线穿过。若三角形一边的两端点为P_1（x_1，y_1，z_1），P_2（x_2，y_2，z_2）则

$(z_1-h)(z_2-h)<0$ 表明该边有等高线点；

$(z_1-h)(z_2-h)>0$ 表明该边无等高线点。

直至搜索到等高线与网边的第一个交点，称该点为搜索起点，也是当前三角形的等高线进入边、线性内插该点的平面坐标（x，y）：

$$\left.\begin{aligned} x &= x_1 + \frac{x_2 - x_1}{z_2 - z_1}(z - z_1) \\ y &= y_1 + \frac{y_2 - y_1}{z_2 - z_1}(z - z_1) \end{aligned}\right\}$$

4）搜索该等高线在该三角形的离去边，也就是相邻三角形的进人边，并内插其平面坐标。搜索与内插方法与上面的搜索起点相同，不同的只是仅对该三角形的另两边作处理。

5）进入相邻三角形，重复第（4）步，直至离去边没有相邻三角形（此时等高线为开曲线）或相邻三角形即搜索起点所在的三角形（此时等高线为闭曲线）时为止。

6）对于开曲线，将已搜索到的等高线点顺序倒过来，并回到搜索起点向另一方向搜索，直至到达边界（即离去边没有相邻三角形）。

7）当一条等高线全部跟踪完后，将其光滑输出，方法与前面所述矩形格网等高线的绘制相同。然后继续三角形的搜索，直至全部三角形处理完，再改变等高线高程，重复以上过程，直到完成全部等高线的绘制为止（图 9-17），图 9-17 描述了利用三角网生成数值为 50 的等高线的过程。

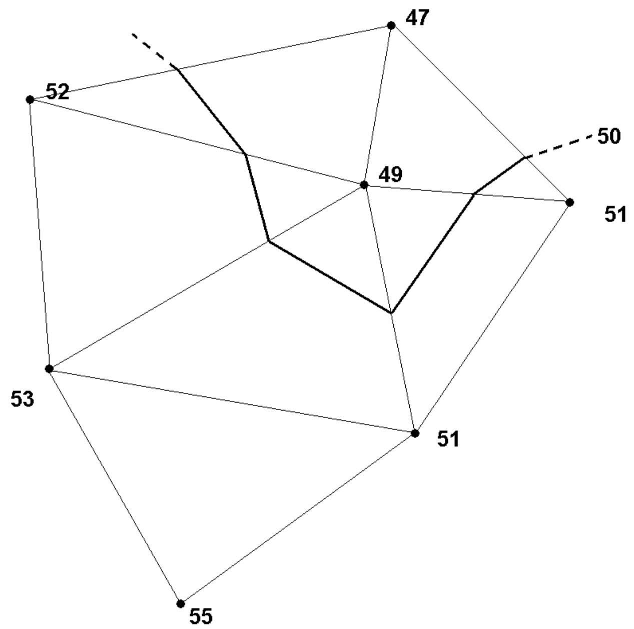


图 9-17：利用 TIN 生成等高线

在其中理论T的所有有效句子都被满足的一种可能的认识，称为T的一个模型。

萨帕斯

一个模型可以是一种理论、一条规律、一种关系、或者一种假说、一个方程式、一条规则。

斯基林

第十章 空间建模与空间决策支持

导读：本章介绍了“更高层次”的GIS分析功能，第一节讲述了如何利用基本的分析，如缓冲区、叠加分析，进行组合，以完成特定的功能。
后面几节分别介绍了空间决策支持系统、专家系统、数据仓库、元胞自动机和空间定位和配置方面的知识，这些内容，或者在其它的领域已经成熟，或者属于新兴的领域。在GIS的支持下，实现相关功能，并将其应用于空间分析，可以解决更为复杂的空间问题。

1. 空间分析过程及其模型

1．1 空间分析过程

空间分析的目的是解决某类与地理空间有关的问题，通常涉及多种空间分析操作的组合。好的空间分析过程设计将十分有利于问题的解决，一般步骤是：

1）明确分析的目的和评价准则；

2）准备分析数据；

3）进行空间分析操作；

4）进行结果分析；

5）解释、评价结果（如有必要，返回步骤1）；

6）结果输出（地图、表格和文档）。

1．1．1 例1：道路拓宽改建过程中的拆迁指标计算

这里将举例说明如何利用建立缓冲区、拓扑叠加和特征提取,计算一条道路拓宽改建过程中的拆迁指标。

1）明确分析的目的和标准

本例的目的是计算由于道路拓宽而需拆迁的建筑物的建筑面积和房产价值,道路拓宽改建的标准是：

- 道路从原有的20m拓宽至60m；
- 拓宽道路应尽量保持直线；
- 部分位于拆迁区内的10层以上的建筑不拆除。

2）准备进行分析的数据

本例需要涉及两类信息，一类是现状道路图；另一类为分析区域内建筑物分布图及相关信息。

3）进行空间操作

首先选择拟拓宽的道路，根据拓宽半径，建立道路的缓冲区。

然后将此缓冲区与建筑物层数据进行拓扑叠加，产生一幅新图，此图包括所有部分或全部位于拓宽区内的建筑物信息。

4）进行统计分析

首先对全部或部分位于拆迁区内的建筑物进行选择，凡部分落入拆迁区且楼层高于 10 层以上的建筑物，将其从选择组中去掉，并对道路的拓宽边界进行局部调整。

然后对所有需拆迁的建筑物进行拆迁指标计算。

5）将分析结果以地图和表格的形式打印输出。

1．1．2 例 2：辅助建设项目选址

本例说明如何利用空间操作和特征提取功能，为一建设项目选择最佳的建设位置。

1）建立分析的目的和标准

分析的目的是确定一些具体的地块，作为一个轻度污染工厂的可能建设位置。工厂选址的标准包括：

- 地块建设用地面积不小于 10000㎡；
- 地块的地价不超过 1 万元/㎡；
- 地块周围不能有幼儿园、学校等公共设施，以免受到工厂生产的影响。

2）从数据库中提取用于选址的数据

为达到选址的目的，需准备两种数据，一种为包括全市所有地块信息的数据层；另一类为全市公共设施（包括幼儿园、学校等）的分布图。

3）进行特征提取和空间拓扑叠加

从地块图中选择所有满足条件 1、2 的地块，并与公共设施层数据进行拓扑叠加。

4）进行邻域分析

对叠加的结果进行邻域分析和特征提取，选择出满足要求的地块。

5）将选择的地块及相关信息以地图和表格形式打印输出。

1．2 空间分析建模

1．2．1 地图模型（Cartographic Model）的概念

模型是人类对事物的一种抽象，人们在正式建造实物前，往往首先建立一个简化的模型，以便抓住问题的要害，剔除与问题无关的非本质的东西，从而使模型比实物更简单明了，易于把握。同样为了解决复杂的空间问题，人们也试图建立一个简化的模型，模拟空间分析过程。空间分析建模，由于是建立在对图层数据的操作上的，又称为“地图建模”（Cartographic Modeling）。它是通过组合空间分析命令操作以回答有关空间现象问题的过程，更形式化一些的定义是通过作用于原始数据和派生数据的一组顺序的、交互的空间分析操作命令，对一个空间决策过程进行的模拟。地图建模的结果得到一个“地图模型”，它是对空间分析过程及其数据的一种图形或符号表示，目的是帮助分析人员组织和规划所要完成的分析过程，并逐步指定完成这一分析过程所需的数据。地图模型也可用于研究说明文档，作为分析研究的参考和素材。

地图建模可以是一个空间分析流程的逆过程，即从分析的最终结果开始，反向一步步分析为得到最终结果，哪些数据是必须的，并确定每一步要输入的数据以及这些数据是如何派生而来。以下的例子将说明其过程：

假定需要获得这样一个结果，即要显示出所有坡度大于20度的地区。首先的问题是要生成这样一幅图像，哪些数据是必须具备的：如要生成一幅坡度大于20度的图像，需要一幅反映所有坡度的图像，数据库里有这样的图像吗？如果没有，就进一步沿着反向思路提问："如要生成一幅所有坡度的图像，需要什么样的数据？"。一幅高程数据图像可用于生成坡度图像。那么，这幅高程数据图像有没有呢？如果没有的话，生成该图像需要何种数据？这一过程一直持续，直至找出所有必备数据为止。然后反向用图形或符号将有关数据及其操作流程表示出来就得到一个地图模型。本例表示如图10-1：

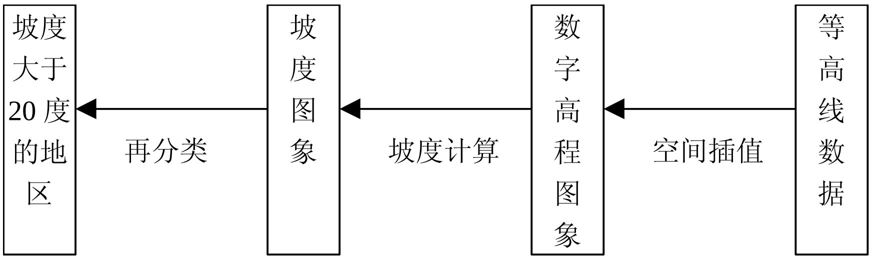


图10-1：提取坡度大于20度的计算流程

图中，矩形框内为数据，箭头表示操作命令，方向表示操作顺序。

1．2．2 地图模型实例

地图模型有多种表示方法，为了进一步理解制图建模过程，下面给出三个不同领域的地图模型实例，分别采用了三种不同的表示方法。

1）食草动物栖息地质量评价模型

本例是一个食草动物栖息地质量评价简化模型，模型只考虑了影响食草动物生存的基本因子：水源、食物、和隐藏条件，以及景观单元的面积，连通性和破碎程度的度量指标。模型形式如图10-2所示：

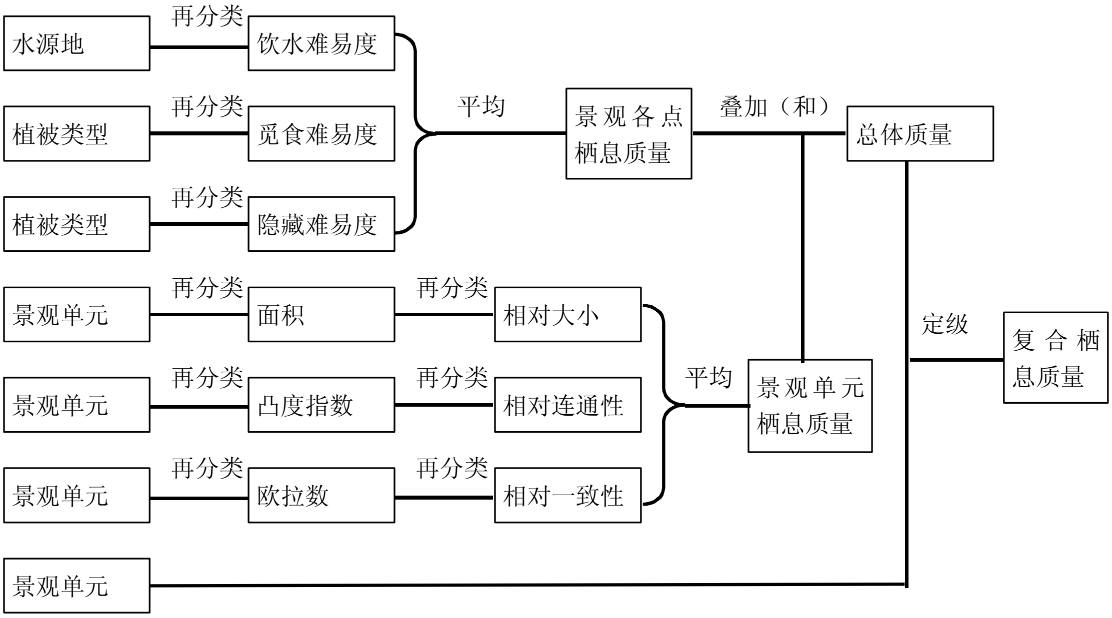


图10-2：食草动物栖息地质量评价模型

图中操作顺序从左向右，从上向下，矩形框内为原始数据和派生中间数据以及结果数据，矩形框连线上面的文字为操作命令。

2）国家森林公园选址模型

本例是一个为某地建立一国家森林公园确定大致范围，是一个数据源已知，需要进行空间信息提取的模型。数据源包括公路铁路分布图（线状地物），森林分布图（面状地物），城镇区划图（面状地物）。地图模型可以用下面的形式表示：

表 10-1：国家森林公园选址模型

步骤	操作命令
找出所有森林地区	再分类
1 为林地 ，0 为非林地	
合并森林分类图属性相同的相邻多边形的边界	归组
找出距公路或铁路 0.5 公里的地区	缓冲区分析
找出距公路或铁路 1 公里的地区	缓冲区分析
找出非城市区用地	再分类
1 为非市区，0 为市区	
找出森林地区、非市区、且距公路或铁路 0.5 至 1 公里范围内的地区	拓扑叠加分析
合并相同属性的多边形	归组

3）木材毁坏量回归预测模型

根据多年的统计数字和经验方程，本例是一个林场砍伐木材时木材毁坏量回归预测模型。

模型的因变量有坡度，树径，树高，蓄积量，树木缺矢量。公式如下：

$$Y = -2.490 + 1.670X_1 + 0.424X_2 - 0.007X_3 - 1.120X_4 - 5.090X_5$$

地图模型可以表示成下面的形式（图10-3）：

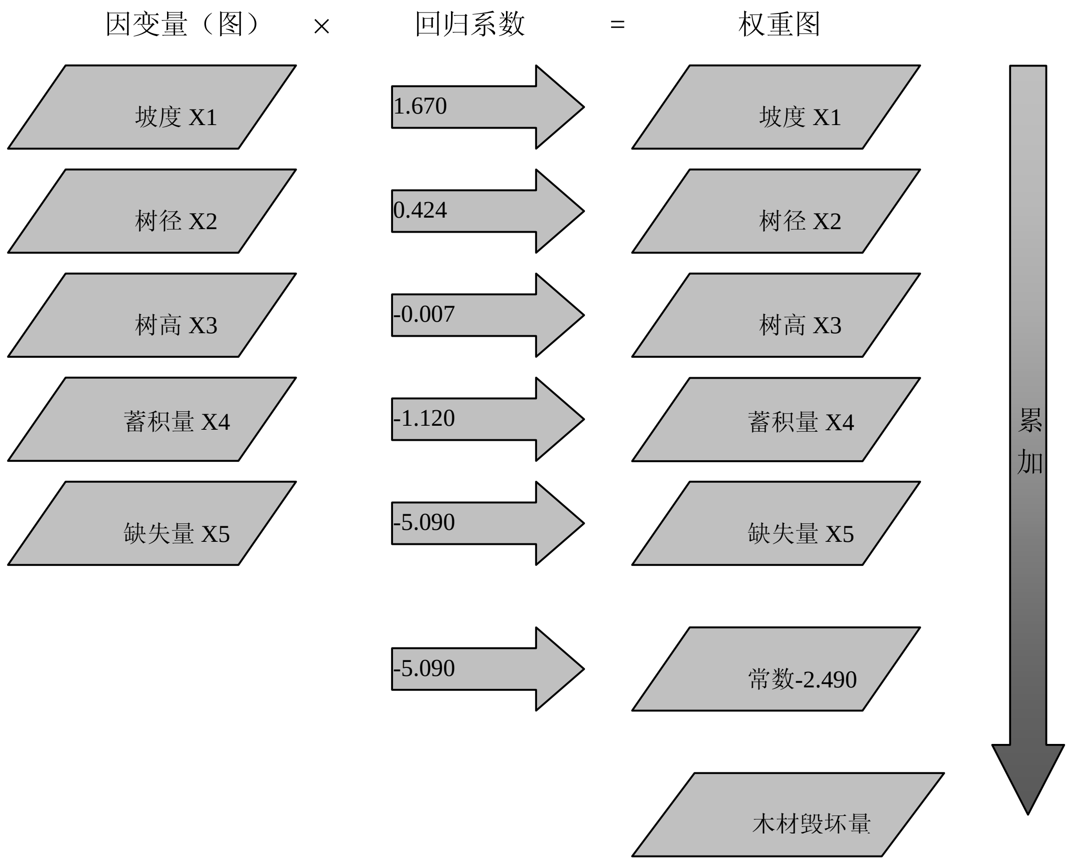


图 10-3：木材毁坏量回归预测模型

1．2．3 地图模型实现

大多数GIS软件提供了宏命令或脚本描述语言，可以将上述建立的各种地图模型表示成GIS的操作命令序列，自动批处理完成整个模型过程。例如一个根据DEM图像生成坡度图，可以表达成GIS命令格式：

CALC　Slopemap = slope (DEMmap)

由多个原始图层生成一个新图可以写成下面的形式：

Newmap = f (Map1,Map2,......)

式中，f ( ) 表示一个GIS命令。

一些GIS软件还提供了书写复杂函数的功能，甚至可以在一个命令行里，使用多个函数表达一个完整的地图模型，形式如下：

Newmap1,Newmap2,......=f1,f2,f3,.......(Map1, Map2,......, Newmap1, Newmap2,......)

式中，Newmap1, Newmap2，... 为派生的中间图层。

还有一些GIS软件提供了高级的可视化的地图建模辅助工具，用户只需使用其提供的工具在窗口中绘出模型的流程图，指定流程图的意义、所用的参数，矩阵等即可完成地图模型的设计，而无需书写复杂的命令程序。可视化地图建模工具为用户提供了高层次的设计工具和手段，可使用户将更多的精力集中于专业领域的研究（图10-4）。

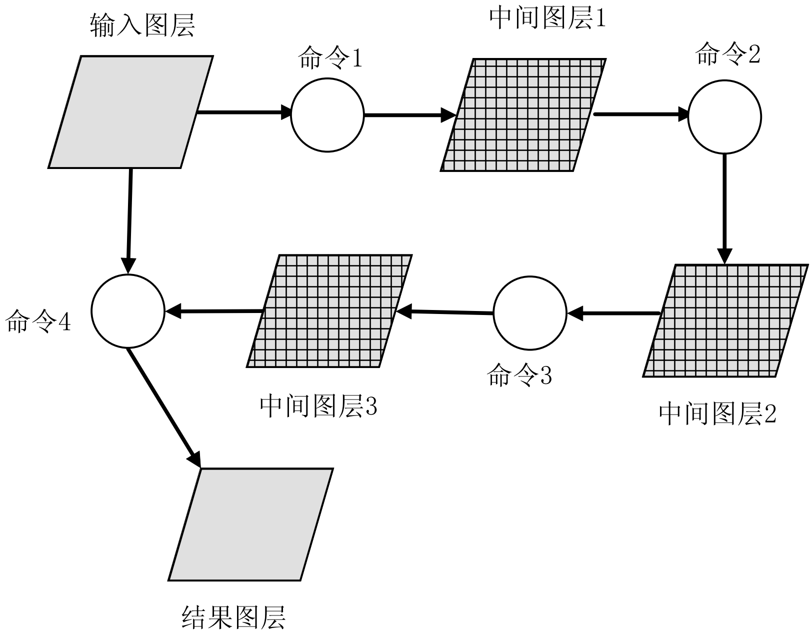


图10-4：通过流程图表现的GIS模型

2．空间决策支持模型

区域规划、土地利用规划、设施位置选择、环境管理等都是有关空间行为的决策问题，这些问题的解决方案是由决策者或领域专家在专业领域知识和经验的启发下，在分析大量的空间和非空间信息的基础上得到的。空间决策问题大大超过了地理信息系统通常的空间分析功能的要求。

2．1 空间决策过程的复杂性

决策是一个决策者为达到某种目标或目标集合，根据一定的约束条件下在多种侯选方案里进行选择的复杂过程。当采用数学表述形式表达一般化决策问题时，包括以下几个构成部分：

1）方案集合：决策问题的方案集合是指可以选择的行动方案集合，记为 A。

2）状态集合：任何一个决策问题都面临一定的外界环境，称之为状态。系统各种可能的状态，称为状态集合，记为 Q。

3）损益函数：这是决策分析中的一个重要概念。在决策问题中，如果采用策略 $a(a \in A)$，假定系统状态出现 $q(q \in Q)$，系统收益 $W = (a,q)$。因此定义映射 $W:(A \times Q) \rightarrow R$ 为决策问题的损益函数。在 A、Q 可数的情况下，可获得损益表如下所示（表 10-2）。

表 10-2：决策损益表

	Q_1	Q_2	...	Q_n

A_1	W_{11}	W_{12}	**…**	W_{1n}
…	**…**	**…**	**…**	**…**
A_m	W_{m1}	W_{m2}	**…**	W_{mn}

4）目标函数（决策准则）：记为 F。损益函数只给出了系统的实际收益情况，但没给出收益的评价标准，即“抉择”时的优化准则。决策准则对于不同的决策者、问题、方法都是不同的，它最终决定了方案的形成。

综上所述，可以将一个决策问题记为：

$$Udm = \{F, A, Q, W\}$$

其中，F 为目标函数或决策准则，A 为侯选方案集，Q 为状态集，W 为损益函数。

决策学常规方法用于解决普通决策问题，这类问题满足以下条件：

- 存在决策者希望达到的明确目标；
- 存在可供决策者选择且可以明确组分的侯选方案；
- 存在不受决策者控制的系统状态，系统状态集与侯选方案集相互独立；
- 损益值可以精确数量化，A，Q 均为可数集合。

当系统状态集 Q 中状态数 n=1 时，为确定性决策问题；当 n>1 时，且系统各状态出现的概率未知时，为不确定性决策问题；当 n>1，且系统各状态出现的概率服从一个已知的概率分布时，为风险决策问题。

空间决策与一般决策问题同样有确定性决策、不确定性决策和风险决策。确定性决策实际上是一个最优化问题，象土地适宜性评价的多准则决策和线性规划均属此类决策问题，它们可以和地理信息系统的空间分析功能完全集成。而大量的空间决策问题往往涉及到结构、非结构化知识，人的评价和判断等不同形式的知识，决策的不确定性和风险性的成分很大。以设施配置为例，领域专家已经有一组有关位置适宜性的判别规则，这些规则属于描述性方式表达的知识，设施位置的选择是建立在有关社会经济、地质条件、环境质量等因素分析的基础上的在判别规则启发下的推理过程；另外领域专家还有有关社会经济、地质条件和环境质量的评价模拟模型，这些知识都属于程式式知识，设施位置的选择是建立在定量模型计算分析的基础上估算过程。

随信息技术的快速发展，为决策者提供了越来越多的空间和非空间的信息，包括地图、航片、表格、遥感和数字测量信号等。决策者需要通过知识和经验来有效的处理和理解这些海量的信息。而人类的知识可分为结构化和非结构化两种知识。结构化的知识有着高度结构化的形式和结构化的求解程序，包括数学模型，统计方法，计算机算法等都属此类型的知识，它们在表现和分析方面遵循固定的框架，大多数情况下只能被专家理解，又称为程式式知识（Procedural Knowledge）。然而大量的知识都是非结构化的，象人类的体验、直觉、价值观，专家经验，本质上是定性的，不能用固定的程序进行表达，又称为描述性知识（Declarative Knowledge）。

决策者使用信息和知识，在解决结构化、非结构化和半结构化问题上的复杂程度大不相同。以设施配置为例，在某些特定约束条件下配置最少数量的设施是一个结构化问题，可以通过最优化方法进行求解；寻找最优设施配置的所有可能的位置则是一个半结构化问题，涉及多种准则评价和价值评判；为设施配置确定总体目标和总体方针政策则属非结构化问题，涉及灵活的定性问题，不能用固定的程式式知识来解决。

总之，空间决策是一个涉及多目标和多约束条件的复杂过程，通常不能简单地通过描述性知识或程式式知识进行解决，往往要求综合地使用信息，领域专家知识和有效地交流手段。空间决策中信息和知识往往是互相作用的，如图 10-5 所示：

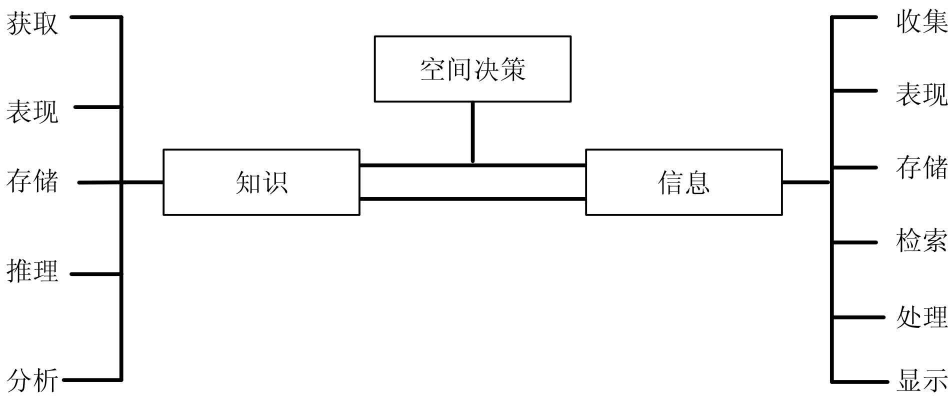


图 10-5：空间决策过程

信息的一侧处理数据的收集、表现、存储、检索、处理和显示，用于计算和量测，以及知识推理和更新。知识的一侧处理知识的获取、表现、存储以及推理和分析，用于处理事实、组织信息和原理。决策中知识和信息的相互作用是对传统信息技术的扩充，没有知识推理不可能做出智能决策。

地理信息系统为决策支持提供了强大的数据输入、存储、检索、显示的工具，但是在分析、模拟和推理方面的功能比较弱，本质上是一个数据丰富但理论贫乏的系统，在解决复杂空间决策问题上缺乏智能推理功能。所以，为解决复杂的空间决策问题，需要在地理信息系统的基础上开发智能决策支持系统，用于数据获取、输入、存储、分析、输出；用于知识表现和推理；用于自动学习，系统集成，人机交互。所用的新技术包括人工智能技术，知识获取、表现、推理等知识工程技术，以及集成数据库、模型、非结构化知识及智能用户界面的软件工程技术。

2．2 空间决策分析的理论和方法

2．2．1 效用理论

效用理论是决策分析的基础。事物的不确定性可看作是许多简单随机事件的复合。每一个简单随机事件是由两个互斥事件Z_1 和Z_2组成的。事件Z_1发生的概率为P，事件Z_2发生的概率为 $1—P$，则随机事件记为L（Z_1，P，Z_2）。在简单随机事件内引进“优先”或 “偏好”的概念，并在随机事件集合的基础上建立公理体系，即假设在随机事件集合中存在下列条件：

1）相对偏好顺序；

2）偏好关系具可传递性；

3）简单随机事件间的可比性；

4）偏好关系可以量化；

5）不确定性可以量化；

6）等价随机事件可相互代换。

在这样的条件下可用一个数值来描述简单随机事件的期望效益，称之为效用。由简单随机事件的效用可确定一般不确定事件的效用。在对事件不确定性判断进行量化时，需要利用各种知识，如系统本身的特性，一些必要的统计知识以及决策者根据经验对事件不确定性的主观估算等。

2．2．2 决策树

决策分析中最常用的方法之一是决策树方法，图 10-6 为典型的决策树。图中长方形小框表示由人选择的决策点。把需要作决策的问题过程画成示意图，由图的最左边出发，在作决策之前先作试验。例如有R个试验L_r，费用为C_r，试验结果有O_1，…，O_t，…，O_T等共T个。在试验L_r条件下O_t 发生的概率记为P_{rt}（O_t）。设此时有d_1，…d_i，…，d_I等共I个备选决策方案。若选择决策di，则这时可能出现S_1，…，S_j，…，S_J共J种状态。在试验L_r中出现结果O_t时选取决策di的条件下，状态S_j出现的概率记为$P_{rti}(S_j)$。此时可能有L种后果x_1，…，x_l; ，…，x_L，而$P_{rti}(S_j)$表示在实验L_r中出现结果为O_t时，选取决策d_i而出现状态S_i的情况下，S_j发生后果x_l的概率，其效用记为U（x_l）。决策树的方法是顺着树的各个分校进行分析，并计算各种可能情况的概率大小，最后计算在这些条件下最终出现的后果的效用，将各种效用加以比较，从中选取最佳效用所对应的试验与决策作为应取的决策。

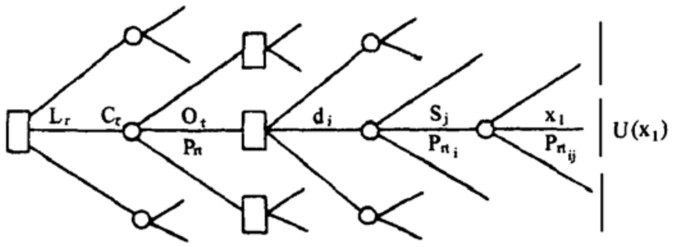


图 10-6：决策树示例

2．2．3 贝叶斯决策

由于决策总是在事件发生之前做出，而事件是否发生又不是确定的，因此常采用统计学中贝叶斯公式对事件发生的概率作先验估计，这就是贝叶斯方法。

由于事件的发生具有不确定性，这就使决策带有一定的经验性。人们对于风验的态度不同，对效用估计也不同。对事件发展持保守看法而不愿冒验的人，对效用估计往往偏低；倾向于冒验的人，对效用的估计往往偏高。也有人取中庸态度，对效用的估计介于两者之间。

2．3 空间决策支持系统

决策支持系统（Decision Support System, DSS）是辅助决策者通过数据、模型、知识以人机交互方式进行半结构化或非结构化决策的计算机应用系统。它是在管理信息系统（Management Information System, MIS）的基础上发展起来的，在 MIS 的基础上增加了非结构化问题处理，模型计算和各种方法，为解决结构化、非结构化和半结构化的问题提供了更广泛的方法。它为决策者提供分析问题、建立模型、模拟决策过程和方案的环境，调用各种信息资源和分析工具，帮助决策者提高决策水平和质量。决策支持系统是辅助管理者对半结构化问题的决策过程，支持而不是代替管理者的判断，提高决策的有效性而不是效率的计算机应用系统。DSS 的基本结构主要由四个部分组成，即数据部分、模型部分、推理机部分、人机交互部分，如图 10-7 所示。

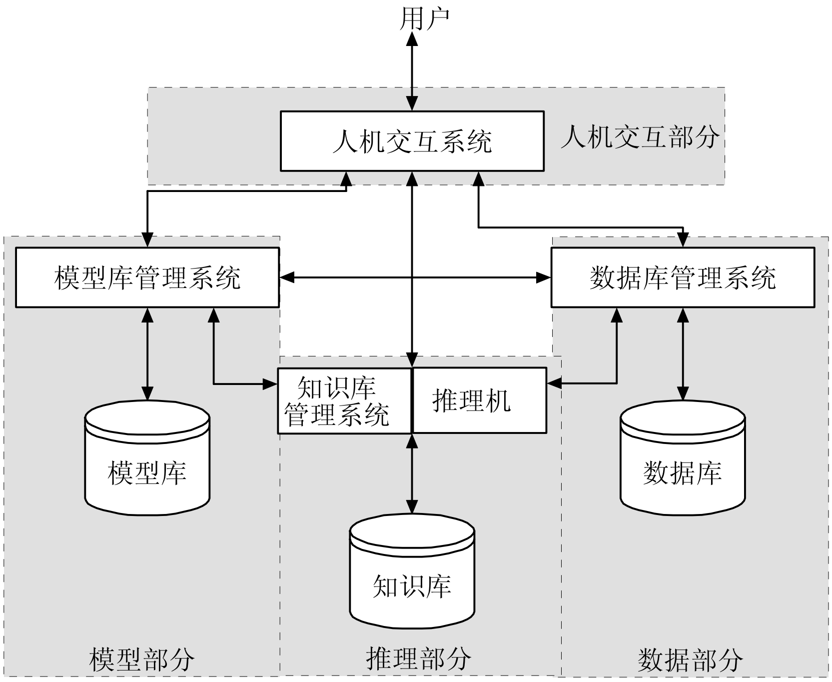


图 10-7：DSS 的组成部分

与 MIS 对应的，GIS 可以看作是用于空间决策的空间信息系统。GIS 与 MIS 的不同之处在于其数据模型和数据结构的复杂性。目前 GIS 的分析功能还很弱且不灵活，它的逻辑结构和智能层次不能满足解决复杂空间决策问题的需要，特别是那些非结构化的问题。为更好地辅助空间决策，GIS 需要增加对描述性知识和程式式知识的处理功能，目前 GIS 还不适合用于对各种知识形式的处理，不能作为空间决策支持系统（SDSS）的神经中枢，但可以作为它的一个组成部分，即 GIS 可以嵌入到一个 SDSS 中，用于空间信息处理。

空间决策支持系统与一般的决策支持系统性质相同，只是更注重空间数据和空间问题的获取和解决。通常空间决策支持系统包括以下的功能：

- 不同数据源的空间和非空间数据的获取、输入、存储；
- 复杂空间数据结构和空间关系表示方法，适于数据查询、检索、分析、显示；
- 灵活的集成程式式空间知识（数学模型、空间统计）和数据的处理功能；
- 灵活的功能修改和扩充机制；
- 友好的人机交互界面；
- 提供决策需要的多种输出；
- 提供非结构化空间知识的形式化表达方法；
- 提供基于领域专家知识的推理机制；
- 提供自动获取知识或自学习功能；
- 提供基于空间信息、描述性知识、程式式知识的智能控制机制。

上述这些功能要求超出了目前 GIS 的范围，需要集成人工智能、知识工程、软件工程、空间信息处理和空间决策理论等领域的最新技术。

2．4 通用智能空间决策支持系统结构体系

空间决策支持系统的建立可以解决特定领域的决策问题。但是它的建立过程是一个花费很长时间的工程，而且它也只能用在特殊的领域，所以建立空间决策支持系统最经济和灵活的方式是使用软件工程和知识工程的方法开发空间决策支持系统开发环境（外壳或产生器），

这样领域专家就可以使用它快速高效地建立多种领域空间决策支持系统。也就是说开发一个通用的开发工具，决策者可以用来定制、修改、调整、扩展空间决策支持系统以解决特定的空间决策问题。图 10-8 是一个通用空间决策支持系统的结构体系图。

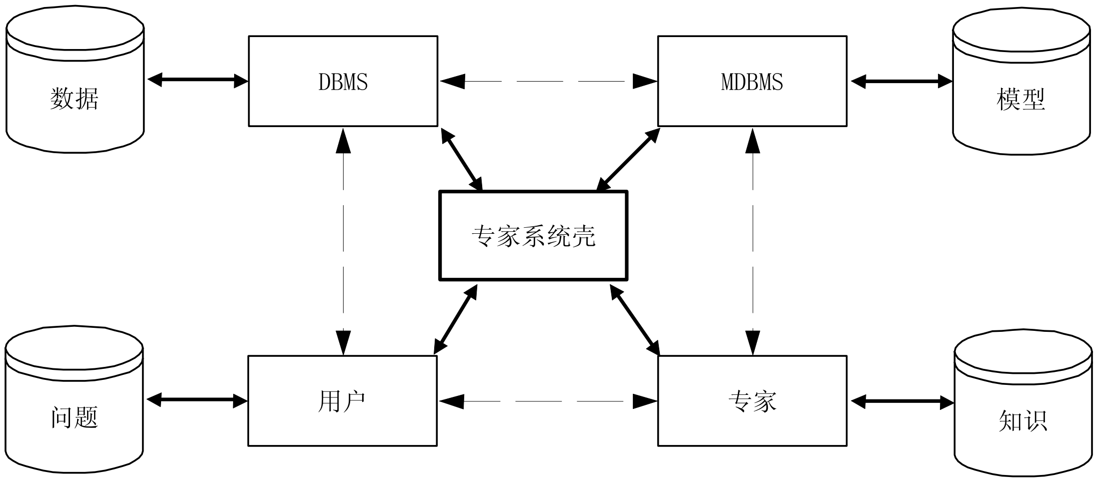


图 10-8：通用空间决策支持系统的结构体系图

系统的核心是一个专家系统壳（Shell），它可以单独作为专家系统开发工具，直接控制着 SDSS 的控制流和信息流，提供表达和存储非结构化领域知识，它还包含了推理控制、系统和用户界面和对外交流的元知识，以及非结构化空间知识的推理机。它是 SDSS 的大脑。为使用空间和非空间数据，专家系统壳有一个与外部数据库的接口，包括 GIS，关系数据库和遥感信息系统。模型管理系统管理和处理程式式知识包括算法、统计程序和数学模型，它也有一个与专家系统壳的接口，可以通过专家系统壳的元知识进行调用。除了与数据库管理系统、模型管理系统的接口外，友好的用户界面和知识获取模块也是专家系统壳的基本组成部分。

本节重点讨论空间决策支持系统的模型管理系统，关于数据库管理系统在前面的章节已作了介绍，至于知识库和知识处理将在下节专家系统里讨论。

2．5 空间决策支持系统的模型管理系统

为解决自然和人文过程中出现的各种复杂的空间问题，多年来学者们提出了大量的结构化模型，包括统计方法、数学模型、启发式程序、算法等，这些模型与描述性的知识不同，具有高度结构化的格式和固定的执行程序。这些模型对解决那些结构化的决策问题很有帮助作用，但不幸的是，它们的形式逻辑和解决方法对决策者来说通常很难或需要花很长时间才能理解，尤其是那些非技术背景的决策者更是容易混淆或不恰当有效地运用它们，从而限制了这种类型的知识的有效利用。另外，这些结构化的模型要在 GIS 环境下使用，还有一个与 GIS 数据模型兼容性的问题，模型与 GIS 数据库相互操作是一个基本要求。所以空间决策支持系统需要适当地挑选和组织有关的模型，与管理空间和非空间数据的数据库管理系统相对应的，要有一个模型管理系统。模型管理系统应具有下列功能：

- 帮助用户选择与分析有关的模型；
- 对多种类型的模型进行分类和维护以支持各种层次的决策过程；
- 能将模型子模块组合复杂的模型；
- 提供恰当的数据结构满足查询、分析、显示；满足与数据库的嵌入或数据交换；满足模型与描述性知识的交流；
- 提供用户咨询和结果解释的友好界面；

空间决策支持系统对模型高效的分类和组织问题是决策支持系统的核心功能。将模型分类并按不同层次的深度进行组织可以有效的管理和使用模型。例如可以先按决策问题进行第一级分类，再按评价条件和状态进行第二级分类，还可以继续进行更深层次的分类。下面给出一个分类示例：

1）决策问题分类

第一级分类：

- 环境问题
- 土地利用规划问题
- 资源分配问题
- 设施配置问题
- 网络问题
- 水文问题
- 地质问题
- 海岸线问题

假设关心的是网络问题，关于网络问题的各种模型组织成第二级分类，如下所示：

- 最短路径
- 最少搜索路径（Spanning tree）问题
- 货郎担问题
- 多点通讯问题（Multicast communication）
- 运输问题
- 商品流问题

对于每个选择的问题，还可以继续细分成更专的类型，如商品流问题，可以继续分成单商品和多商品流问题。要选择某个专用的模型，用户通过一系列“是或否”的问题向导，直到找到需要的解决问题的模型。

2）按技术条件分类

表 10-3：决策问题的分类示例

A．确定性 确定性模型 不确定性模型 　　随机模型 　　不精确模型	B．空间 离散模型 连续模型
C．过程 静态模型 动态模型	D．时间 离散模型 连续模型
E．线性 线性模型 非线性模型	F．目标 单目标模型 多目标模型
G．变量 实型 整型	

根据上述分类，可以构造一个决策树，将模型分类知识用一种知识表达方式进行表示，例如本例可以用产生式规则表示，每条路径对应一个规则。本例对确定性分类下共有 2^6规则，其中一条规则描述如下：

表 10-4：决策规则

IF situition certain　　AND

space discrete AND
process static AND
time discrete AND
system linear AND
objectives multiple AND
variables real
THEN select multiple-objectives, discrete-space,
discrete-time linear programming model

IF situition uncertain AND
cause random AND
space discrete AND
process dynamic AND
time discrete AND
system linear AND
objectives single AND
variables real
THEN select multiple-stage, single-objective, discrete-space,
discrete-time linear stochastic programming model

空间决策支持系统除了模型选择的问题外，模型与数据库的交互也是一个重要的问题。不同的模型具有不同的数据结构和不同的模型与数据库的交互方式。模型要在 GIS 环境下运行，还有一个与 GIS 数据结构的兼容性问题。

模型与 GIS 有不同层次的交互方式。最低层的交互是将 GIS 作为数据库管理系统，以文件的方式进行交互，如果模型和 GIS 对文件类型是兼容的，交互只是一个简单地文件选取问题；如果文件不兼容，就会涉及文件转换的问题。较高层次的交互方式是将 GIS 作为显示和分析模型结果的图形显示工具。最高层次的交互是二者集成在一个完整的系统里，比较松散一点的方式是用 GIS 的操作命令（如宏语言）重新实现模型；更紧密的集成是二者具有支持查询、分析、显示的数据结构。由于模型的多样性，数据结构也是多样的，因此很难作到 GIS 数据结构与所有模型的结构兼容。所以模型与 GIS 的交互在空间决策支持系统里允许有多种方式。

3．专家系统

3．1 专家系统的基本组成

专家系统是人工智能在信息系统中的应用，它是一个智能计算机程序系统，其内部具有大量专家水平的某个领域知识与经验，能够利用人类专家的知识和解决问题的方法来解决该领域的问题。专家系统的主要功能取决于大量的知识。设计专家系统的关键是知识表达和知识运用。专家系统与一般计算机程序最本质的区别在于：专家系统所解决的问题一般没有算法解，并且往往是要在不完全、不精确或不确定的信息基础上做出结论。一般的专家系统包括数据库，知识库，推理机，解释器及知识获取五个部分组成，它的结构如图 10-9 所示。

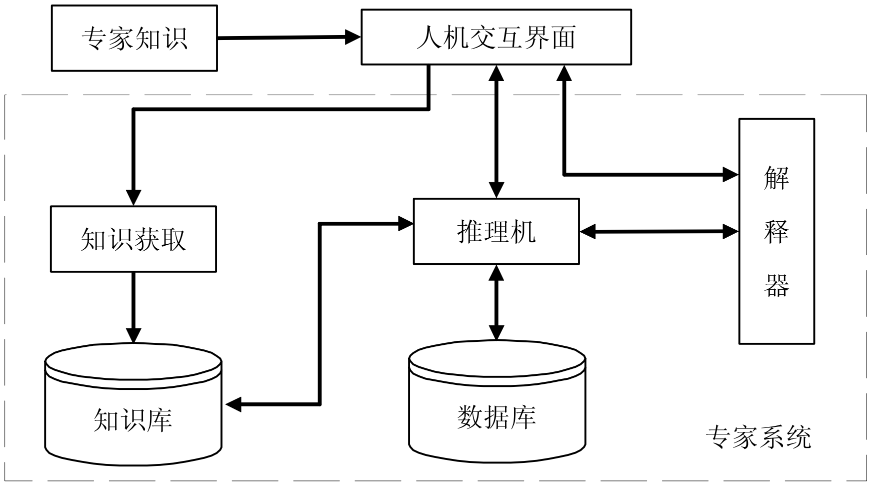


图 10-9：专家系统结构图

1）知识库

知识库用于存取和管理所获取的专家知识和经验，供推理机利用，具有知识存储、检索、编辑、增删、修改和扩充功能。

2）数据库

用来存放系统推理过程中用到的控制信息，中间假设和中间结果。

3）推理机

用于利用知识进行推理，求解专门问题，具有启发推理、算法推理；正向、反向或双向推理；串行或并行推理等功能。

4）解释器

解释器用于作为专家系统与用户的“人——机”接口，其功能是向用户解释系统的行为，包括：

（4．1） 咨询理解：对用户咨询的提问进行“理解”，将用户输入的提问及有关事实、数据和条件、转换为推理机可接收的信息。

（4．2） 结论解释：向用户输出推理的结论或答案，并且根据用户需要对推理过程进行解释，给出结论的可信度估计。

5）知识获取器

知识获取是专家系统与专家的“界面”。知识库中的知识一般都是通过“人工移植”方法获得，“界面”就是知识工程师（专家系统的设计者），采用“专题面谈”，“口语记录分析”等方式获取知识，经过整理后，再输入知识库。为了提高知识工程师获得专家知识的效率，可以借助“知识获取辅助工具”来辅助专家整理知识或辅助扩充和修改数据库。近年来，开始机器学习、机器识别、半自动化等方法获取知识。

3．2 专家系统的知识处理

3．2．1 人工智能与专家系统

人工智能的目的是用计算机模拟人类，其中包括模拟人类的动作——机器人，模拟人类的视听能力——计算机视觉、听觉的模式识别，模拟人类语言——计算机自然语言，模拟人脑——电脑。在模拟人脑研究时，本来是想从幼儿开始，随着大脑的发育过程模拟人脑的，

但是发现这样模拟难度太大了，转而研究领域专家的思维过程，因为领域专家能够阐述清楚知识与逻辑。随着专家系统工具软件的出现，许多领域展开了专家系统的研究，其中最成功的是国际象棋的专家系统。人工智能主要的目的是模拟人脑的功能，但是目前人们对人脑的思维过程并不十分清楚，因此人工智能的概念也不可能非常清楚。许多人工智能的研究只局限于形式逻辑的推导，凡是超出了形式逻辑范畴的，都被是认为无法解决的问题。现在理解的人工智能主要是指用计算机完成逻辑推理的过程。

3．2．2 知识表示

知识表示就是知识的形式化，就是研究用机器表示知识的可行的、有效的、通用的原则和方法。目前常用的知识表示方法有：产生式规则、语义网络法、框架表示法、与或图法，过程表示法、特征表示法、黑板结构、Petri 网络法、神经网络等。其中效果最好的是产生式规则，其它表示方法单独使用的不多，大多是以产生式规则为主体增加和扩展语义、框架和程式式知识。

产生式规则使用“若……，则……。”语句（If…,Then)。“若”可以是一个条件，也可以是多个条件；“则”也可以是一个结论，也可以是多个结论。也称为正向式推理，有以下几种情况：

1）一对一：*If (1), then (a)*。

2）一对多：*If (1), then (a), (b),… (n)*。

3）多对一：*If (1), (2), …,(m); then (a)*。

4）多对多：*If (1), (2), …,(m); then (a),(b),…,(n)*。

If…,then 语句是从已知条件，推论结果的语句，反之，从结果推出条件称为后向式推理。同样有四种情况，一对一，一对多，多对一，多对多：

1）一对一：*Conclusion (a);Need(1)*。

2）一对多：*Conclusion (a);Need(1),(2),…,(m)*。

3）多对一：*Conclusion (a),(b),…,(n);Need (1)*。

4）多对多：*Conclusion (a),(b),…,(n);Need (1),(2),…,(m)*。

由此可见，产生式规则的正向式推理与逆向式推理都是在确定性问题中的处理方法。对于不确定性的问题，产生式规则不能解决问题，需要结合其它的表示方法。

3．2．3 知识推理

推理是指依据一定规则从已有的事实推出结论的过程。专家系统中的自动推理是知识推理，它是专家系统中问题求解的主要手段。

知识推理与知识表示有密切关系，根据知识表示的特点，知识推理方法可分为图搜索方法和逻辑论证方法两类。

．**图搜索方法**

在专家系统的知识表示中，许多基本的、常用的表达方式都具有“图”的形式，或者可以变换为相应的图的形式，而且通常可以用与或图来进行表达。例如，状态空间图、与或图、语义网格图，以及由产生式规则或框架表示方法所转换的与或图或网络图。

基于图的知识表达，问题求解的知识推理过程，就是从图中相当于初始状态的出发节点到相当于目标状态的终止节点的路线搜索过程，即搜索从初始状态有效的转移到目标所经历的最优的或最经济的线路，相应的知识推理方法即图搜索方法。例如，对于具有树状的状态空间图，称为“问题树”，基本的图搜索方法有：宽度优先搜索、深度优先搜索法等。

．逻辑论证方法

当知识表示采用谓词（Predicate）逻辑或其他形式逻辑方法时，知识推理也可以采取逻辑论证方法。在这种情况下，求解一个问题相应于证明一个定理或几个定理，问题求解的知识推理过程，相应于用数理逻辑方法进行定理证明的过程。知识推理方法即逻辑论证方法。例如，若用一组谓词逻辑表达式 A 描述有关的事实、情况或条件，而用另一组谓词逻辑表达式 B 描述问题的答案或结论，那么，只要通过逻辑演算方法论证定理 A-B 成立，也就相应论证了从有关事实、情况和条件出发，可以推出正确的答案或结论 B。

根据问题求解的推理过程是否运用启发性知识，知识推理方法可分为启发推理和非启发推理两类。

．启发推理

在问题求解的过程中，运用与问题有关的启发性知识，即解决问题的策略、技巧，对解的特性及其规律的估计等实践经验或知识，以加快推理过程，提高搜索效率，这种推理过程称为“启发式推理”。例如，在图搜索的推理方法中，利用启发性知识改进的深度优先搜索法，如局部择优搜索法，最好优先搜索法等，只需要对部分状态空间进行搜索，可提高搜索效率。

．非启发推理

在问题求解的推理过程中，不运用启发性知识，只按照一般的逻辑法则和控制性知识，进行通用性的推理。这种方法缺乏对求解问题的针对性，需要进行全状态空间的搜索，而没有选择最优的搜索途径，大多搜索效率低。例如宽度优先搜索法，虽然是完备的算法，但其搜索效率低。

根据问题求解的推理过程中结论是否精确，知识推理方法可分为精确推理和不精确推理两类。

．精确推理 精确推理是指在专家系统中，把特定领域的知识表示成必然的因果关系、逻辑关系，推理的结论是肯定的。这种推理是精确推理。

．不精确推理 在人类知识中，有相当一类属于人们的主观判断，是不精确和含糊的。由这些知识归纳出来的推理规则也往往是不确定的，基于这种不确定的推理规则进行推理，形成结论，称为不精确推理。常用的不精确推理方法有概率论方法、可信度方法、模糊子集法和证据论方法。

根据问题求解过程中特殊和一般的关系，知识推理方法可分为演绎推理和归纳推理；根据求解推理过程中的推理的方向，知识推理方法可分为正向推理、反向推理和正反向混合推理三类。

3．3 空间分类专家系统实例——土地类型分类

空间分类是 GIS 和遥感信息系统最常用的功能，将空间单元归组分类是智能 GIS 的基本功能。传统的分类基于二值逻辑，认为区域分类界限是明显的，基本空间单元（矢量或栅格结构）属于且只属于一个空间类别。然而，事实上空间类别之间的界限通常是模糊的，是渐变的而不是突变的，所以在空间分类中引入模糊逻辑会提高分类的精度。另外空间分类也是一个基于人们对空间现象的认知和知识的心理判断过程，开发一个带有 GIS 的专家系统是很有必要的。

当然可以使用那些结构化的分类算法（程式式知识），但是这种方法往往是很机械，且不能有效地与用户进行关于分类的知识相互交流。专家系统的分类方法则更为灵活、智能，它不是一个算法，而是一个关于空间分类的抽象和空间数据如何分类的规则，可以在不修改

任何程序的情况下在知识库里修改、删除、增加专家和非专家的判断，分类知识和规则。基于规则的专家分类系统往往具有传播知识和教育的成分。下面是一个基于遥感影象的土地类型专家分类系统实例。

数据采用 4 波段 LANDSAT MSS 影象数据，分类系统所采用的分类依据如图 10-10 所示：

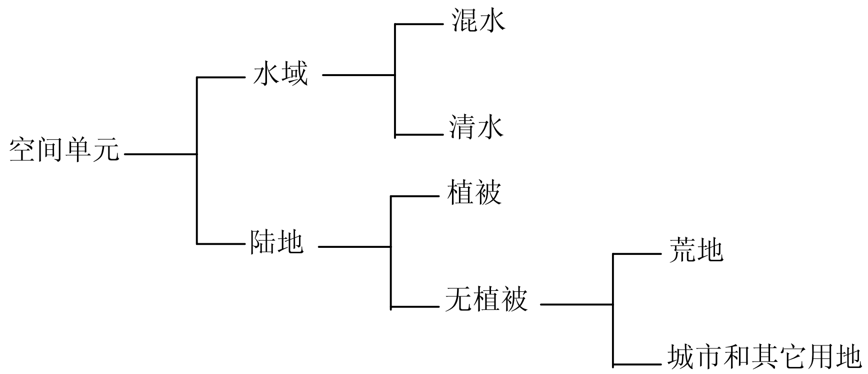


图 10-10：土地利用分类

由于分类问题是一个数据驱动的过程，采用了正向推理方式。产生式规则是从领域专家获取的，如下表所示，存入知识库。

表 10-5：知识库存储的分类规则

```
rule r1
        If（x1<8.0000,12.0000）
          and (x4<5.0000,10.0000)
          then pretype1 is water
        Certainty is 1
rule r2
        If ( x3>=8.0000,12.0000)
          and (x4>=5.0000,10.0000 )
        then pretype1 is land
rule r3
        If ( pretype1 is water )
          and    (x1>20.0000,24.0000）
          and    (x2>13.0000,16.0000)
          and    (x3>5.0000,10.0000)
          then type is turbid
        Certainty is 1
rule r4
        If ( pretype1 is water)
          and ( x1<=20.0000,24.0000)
          or    (x2<13.0000,16.0000 )
          or    (x3<=5.0000,10.0000 )
        then type is clear
        Certainty is 1
rule r5
        If ( pretype1 is land )
          and    (x1<17.0000,25.0000）
          and    (x2<13.0000,30.0000)
          and    (x3>27.0000,50.0000)
          and    (x4>30.0000,65.0000)
        then type is vegetation
        Certainty is 1
rule r6
```

```
        If ( pretype1 is land)
          and ( x1>=17.0000,25.0000)
          or    (x2>=13.0000,30.0000 )
          or    (x3<=27.0000,50.0000 )
          or    (x4<=30.0000,65.0000 )
        then pretype2 is non-vegetation
        Certainty is 1
rule r7
        If ( pretype2 is non-vegetation )
          and    (x1>27.0000,39.0000）
          and    (x2<26.0000,52.0000)
          and    (x3>26.0000,49.0000)
          and    (x4>21.0000,41.0000)
        then type is barren
        Certainty is 1
rule r8
        If ( pretype2 is non-vegetation)
          and ( x1<20.0000,30.0000)
          and    (x2<20.0000,30.0000)
          and    (x3>20.0000,35.0000)
          and    (x4>15.0000,25.0000)
        then type is urban_or_other
        Certainty is 1
```

根据专家的经验，不能武断的使用一个波段数值作为分类的界限，为反映光谱反射的渐变性，使用了模糊逻辑方法表示推理规则。以规则 r1 为例，在划分水域时不是只取波段 3（x3）的一个数值 5 和波段 4（x4）中的一个数值 8 作为分类依据，而是用了模糊集的方法允许渐变特征。模糊集表示如图 10-11。

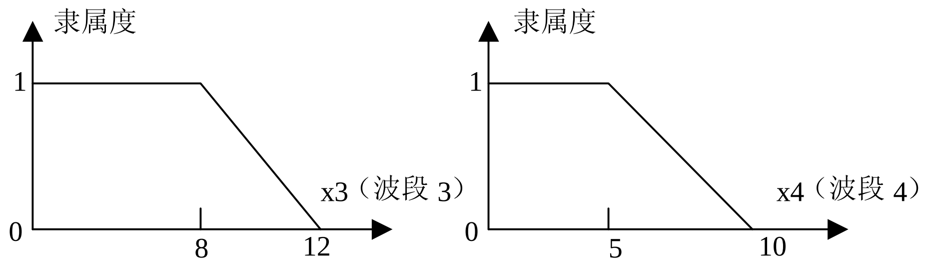


图 10-11：模糊集

由于专家对分类规则非常确定，所以确定性因子为 1。若确定性水平的不同，确定性因子可在[0，1]范围内取值。

分类的结果通过地理信息系统进行显示，当分类结果精确时，系统会自动显示分类结果。当分类结果存在误差时，混合象元同时被标识和显示，同时土地类型渐变造成的模糊边界也可显示，模糊边界的宽度在土地类型混杂的地方用不同的阀值控制。

4. 数据仓库与空间数据挖掘

随着卫星和遥感技术以及其它自动数据采集工具的越来越广泛的应用，日益丰富的空间和非空间数据收集和存储于大空间数据库中，海量的地理数据在一定程度上已经超过了人们能够处理的能力，从这些海量的数据中提取和发现地学知识，给当前 GIS 技术提出了巨大的挑战。新的需求推动着 GIS 从操作型信息系统向分析型信息系统过渡，数据库系统最新的数据仓库和数据库知识发现技术为 GIS 组织、管理海量空间和非空间数据提供了新的思

路。GIS 在空间数据存储、表达和管理方面的能力已得到广泛的认可，增强 GIS 分析功能，提高 GIS 解决地学实际问题的能力已得到共识。GIS 吸收数据仓库的思想，将空间分析和空间数据挖掘方法紧密集成，充分利用 GIS 数据存储、管理空间数据的功能，使海量的地理空间数据变成无限的知识，使 GIS 成为智能的信息系统。

4．1 数据仓库 （Data Warehouse）

近年来，人们逐渐认识到计算机系统存在两类不同的处理：操作型处理和分析型处理。操作型处理也叫事务型处理，是指对数据库联机进行的日常操作，通常是对一个或一组记录的查询和修改，主要为企业特定应用服务，对此人们关心的是响应时间、数据的安全性和完整性。分析型处理是用于管理人员的决策分析，例如决策支持系统、专家系统和多维分析等，经常要访问大量的历史数据。

两者的巨大差别使得操作型处理和分析型处理的分离成为必然，于是数据库由操作型环境发展成为操作型环境和分析型环境的新体系化环境。在此新体系环境中数据仓库处于核心地位，它是建立决策支持系统的基础。

数据仓库是指面向主题的、集成的、稳定的、随着时间变化的数据集合，用以支持管理决策。这一定义指出了数据仓库的目标是为了制定管理的决策提供支持信息。正象企业为了发展要进行业务重组一样，为了支持管理决策需要也要按决策业务科目的要求重组在线事务处理（OLTP）系统中的数据，并要按不同决策，分析内容分别组织使之方便使用。也就是说，数据仓库是一种把收集的数据转变成有意义的信息技术。数据可以来源于许多不同的数据源，包括不同的数据库系统，甚至来源于不同的操作系统。

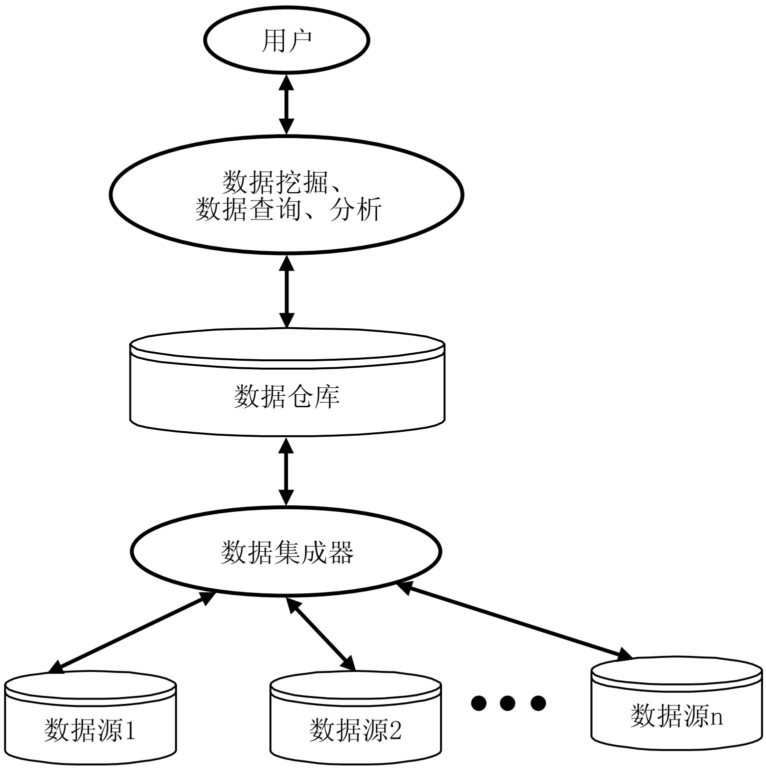


图 10-12：数据仓库体系结构

在把数据装载到数据结构重组后的数据仓库之前，先要进行数据集成处理。这一处理包括几个必不可少的操作步骤，以做到使数据完整、统一，这就确保了在使用数据仓库时其中

的数据是有质量保证的。简而言之，集成就是保证数据准确、到位、没有超出应有的数值范围、没有重复等。数据仓库中的数据不像在线事务处理系统中的数据那样频繁的修改，所以它是比较稳定的（极少或根本不修改）。在一次数据分析的执行过程中使用的数据不得变更，才能保证两次在使用同一组信息进行分析时不会得出不同的答案。数据仓库一般是按周、月或隔月从在线事务处理系统周期性地更新数据。

典型的数据仓库体系结构如图 10-12，数据从多个操作型数据库和外部文件中抽取，抽取出来的数据要进行清理、转换和集成，然后装入数据仓库中。装入仓库的数据形式取决于数据仓库里数据库的设计，一般的数据仓库设计方法是多维数据模型，具体表现为星形模式或雪花模式。仓库的数据要定期更新以反映源数据的变化。最后使用前端的报表、查询、分析和数据挖掘等工具来操作和使用数据仓库。在数据仓库管理方面，有一个存储元数据（关于仓库数据的数据）的中央数据库或数据字典，而监控和管理工具也是必不可少的。

4．2 数据挖掘（Data Mining）

随着大量的大规模的数据库迅速不断地增长，人们对数据库的应用已不满足于仅对数据库进行查询和检索。仅用查询检索不能帮助用户从数据中提取带有结论性的有用信息。这样数据库中蕴藏的丰富知识，就得不到充分的发掘和利用，形成“数据丰富而知识贫乏”的现象。另外，从人工智能应用来看，专家系统的研究虽然取得了一定的进展。但是，知识获取仍然是专家系统研究中的瓶颈。知识工程师从领域专家处获取知识是非常复杂的个人到个人之间的交互过程，具有很强的个性，没有统一的办法。因此，有必要考虑从数据库中发现新的知识，被称为数据库知识发现（Knowledge Discovery in Databases，简称 KDD)，也叫数据挖掘（Data Mining）。数据库知识发现或数据挖掘的定义为从数据中提取隐含的、先前不知道的和潜在有用的知识的过程。数据挖掘技术集成了机器学习、数据库系统、数据可视化、统计和信息理论等多领域的最新技术，有着广泛的应用前景。

数据挖掘主要分为以下四个步骤：

1）数据选取

数据仓库中的数据并不都与挖掘的信息有关，第一步就是为了只提取“有用的”数据。

2）数据转换

在确定要进行挖掘的数据之后，要对这些数据进行必要的变换，使得数据可以被进一步的操作使用，通常的变换有：

将定名量转换为定序量，以便于人工神经网络运算；

对已有的属性进行数学或逻辑运算，以创建新的属性。

3）数据挖掘

在数据转换之后，就要进行数据挖掘，数据挖掘的具体技术很多，如分类、回归分析等。

4）结果解释

挖掘的信息要参照用户的决策支持目的进行分析，并且要表现给决策者。这样，结果的输出不仅包含可视化的过程，而且要经过过滤，以去掉决策者不关心的内容。

当执行完一个挖掘过程后，有时可能需要重新修改挖掘过程，还可能增加其它数据，数据挖掘过程可以通过适当的反馈反复进行，如图 10-13 所示。

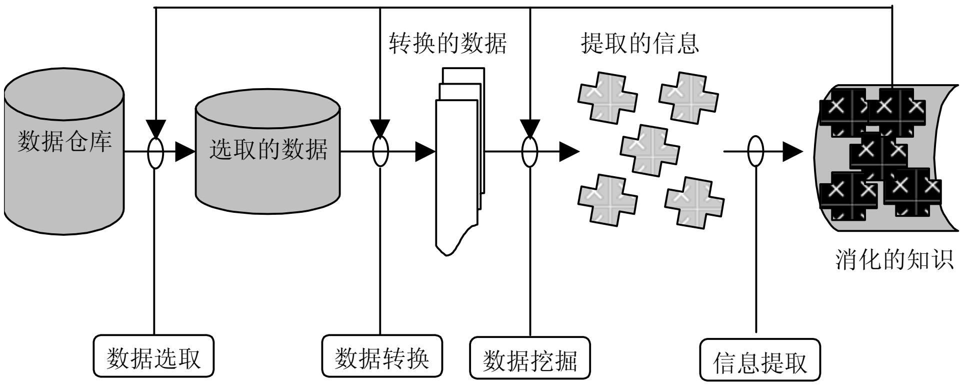


图 10-13：知识挖掘过程

数据挖掘涉及的学科领域和方法很多，有多种分类法。根据知识发现任务，可分为分类或预测模型发现、数据总结、聚类、关联规则发现、序列模式发现、依赖关系或依赖模型发现、异常和趋势发现等等；根据知识发现对象，可分为关系数据库、面向对象数据库、空间数据库、时间数据库、文本数据源、多媒体数据库、异质数据库、Web 数据库；根据知识发现方法，可粗分为机器学习方法、统计方法、神经网络方法和数据库方法。机器学习中，可细分为归纳学习方法(决策树、规则归纳等)、基于范例学习、遗传算法等。统计方法中，可细分为回归分析(多元回归、自回归等)、判别分析(贝叶斯判别、费歇尔判别、非参数判别等)、聚类分析(系统聚类、动态聚类等)、探索性分析(主元分析法、相关分析法等)等。神经网络方法中，可细分为前向神经网络(BP 算法等)、自组织神经网络(自组织特征映射、竞争学习等)等。数据库方法主要是多维数据分析或在线事务处理方法，另外还有面向属性的归纳方法。

4．3 空间数据挖掘

空间数据是与占有一定空间的对象有关的数据，空间数据库是通过空间数据类型和空间关系存储和管理空间数据。空间数据通常具有拓扑和距离信息，通过空间索引进行组织和查询。空间数据特有的性质给空间数据库的知识发现提出了挑战和机遇。空间数据库的知识发现或空间数据挖掘，可以定义为从空间数据库中提取隐含的知识、和没有直接存储的空间关系、空间模式的过程。

空间数据挖掘技术，特别是空间数据理解、空间和非空间数据关系发现、空间知识库构造、空间数据库的查询优化和数据组织，在 GIS、遥感、影象数据库、机器人运动等涉及空间数据的应用系统中很有前景。

下面简要介绍目前空间数据挖掘使用的一些方法：

1）统计分析方法

目前空间数据分析最常用的方法是统计分析方法。统计分析非常适合处理数值型数据，统计分析方法多年来积累了大量的算法，可以用于对空间现象建模和分析。但空间统计分析也有很大的缺陷：首先统计分析方法关于空间分布数据的统计独立性的假设，往往不现实，因为空间邻域之间存在着相互关系；其次，统计分析不适合处理非数值型数据，如空间对象的名称和定名数据类型等。另外统计分析往往对领域专家和统计方面的知识要求较高，只适合领域专家和有统计经验的人使用，而且当数据不完整或不充分时，统计分析的结果缺乏实际意义。再就是统计分析的计算代价也是很高的。为克服统计分析方法的缺点，需要新的数据挖掘方法。

2）基于概括的方法

基于概括的方法是一种面向属性的归纳学习方法，用于空间数据挖掘，可将空间和非空间属性的关系概括成高层次的概念知识。它需要背景知识，即概念层次体系（Concept Hierarchy），常以概念树的形式给出。如图 10-14 所示的是一个农业土地利用的概念树。同样地，空间数据也存在类似地概念层次，如乡镇——县市——省——国家。

面向属性的归纳学习方法，通过概念树攀升而得到概括的数据，概括的数据可直接转换成规则或逻辑表达式，可用于发现普遍特征规则和区划规则。空间数据库的归纳学习可以根据非空间数据的概念层次，也可根据描述空间数据的概念层次进行。基于非空间数据的概念层次归纳，例如数值型数据可以归到一个数值范围或一个较高层次概念描述型变量，（-9℃在-10℃-0℃范围，或“寒冷”），通过低层次的数值不同的数据概括为较高层次的相同的数据，这些由不同元组的数据和它们的空间对象指针一起合并为高层次数据，表示一个新的空间属性。基于空间数据的概念层次归纳，可以通过区划划分或空间存储结构如四叉树进行定义。

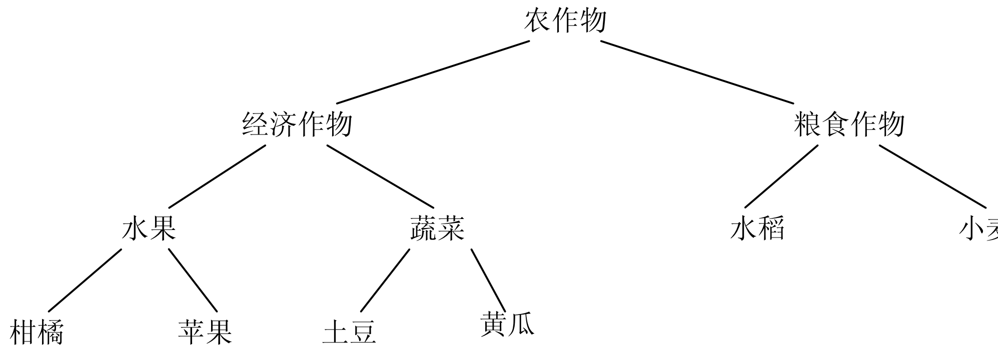


图 10-14：概括归纳示例

基于概括的方法一般有以非空间数据为主的概括和以空间数据为主的概括两种算法。这两种算法的第一步都是选取用户特定的查询数据。接下来，非空间数据为主的概括算法过程是：（A）概念树攀升，即将元组的属性变为高层次的属性；（B）属性删除，删除那些不可能概括的特殊属性值；（C）合并标识后的元组。归纳过程直到所有的属性概括到合适的层次为止。最后将概括后属性值相同的相邻区域合并，结果得到一幅概念层次较高的、而区域较少的图。

以空间数据为主的概括算法，第一步也是收集用户查询的数据，然后根据给定的空间数据概念层次对空间对象进行合并，概括过程直到达到需要的概念层次为止。然后对于每个概括后的空间对象进行非空间数据概括，直到得到所有区域的正确描述。

3）聚类方法

聚类分析是统计分析的一个分支，这种方法的主要优点是不需要背景知识，可以直接从数据中发现感兴趣的结构或聚类模式，类似于机器学习中的非监督学习。

聚类分析与面向属性归纳推理结合，为描述相似对象的空间行为提供了可能，或用于判定不同类别的特征。同样可分为以空间数据为主和以非空间数据为主两种方法。以空间数据为主的方法，首先通过基于采样的高效的聚类算法将与任务有关的空间对象（如点）分类，接着对非空间数据使用面向属性归纳推理方法，提取描述各类一般属性。

以非空间数据为主的方法，首先将与任务有关的空间对象概括到较高的概念层次，进行聚类分析，然后将空间对象进行合并处理。

4）空间关联规则（Spatial Association Rule）方法

关联规则的提出源于交易数据集中发现商品之间隐含的相互依赖关系。如果在超市的数据库中发现的一条规则，形式为“W→B（C%）”，规则的含义是“如果交易中出现了 W 模式，就有 C%的可能性（可信度）出现 B 交易模式”。例如规则“牛奶→黄油（90%）”，表明

有 90%的买牛奶的顾客会同时买黄油，这就是关联规则。它同样可用于发现空间关系。

空间关联规则的形式为：

$$P_1 \wedge P_2 \cdots \wedge P_m \rightarrow Q_1 \wedge Q_2 \cdots \wedge Q_n (c\%)$$

谓词 $P_1, P_2 \cdots P_m, Q_1, Q_2 \cdots Q_n$ 中至少有一个是空间谓词；C%为关联规则的可信度。

构造空间关联规则的空间谓词有很多形式，例如拓扑关系（相交、包含、重叠等）、方位关系（左侧、西侧等）、距离关系（临近、远离等）。

关联规则通常可以分为两种：布尔型关联规则和多值关联规则。多值关联规则比较复杂，一种自然的想法是将它转换为布尔型关联规则。当全部属性的取值数量都是有限的时候，只需将每个属性值映射为一个布尔型属性即可。当属性的取值范围很宽时，则需将其分为若干区段，然后将每个区段映射为一个布尔型属性。于是，如何划分区段是实现多值关联规则到布尔型关联规则转变的关键。这里面有两个互相牵制的问题：当区段的范围太窄时，则可能使每个区段对应的属性的支持度很低，而出现“最小支持度问题”；当区段的范围太宽时，则可能使每个区段对应的属性的可信度很低，而出现“最小可信度问题”。

在一个集合 S 里，空间联结谓词 $P = P_1 \wedge P_2 \cdots \wedge P_k$ 的支持度 $\sigma(P/S)$，为满足 P 的对象数与 S 集合总数的比；集合 S 里关联规则 P→Q 的可信度 $\varphi(P \wedge Q/S)$ 为 $\varphi(P \wedge Q/S) / \sigma(P/S)$。

关联规则的发现可以分解成以下两个子问题：

（4．1）找出数据库 S 中所有满足用户指定最小支持度的对象集 P（一个非空子集），具有最小支持度的对象集称为频繁对象集，反之就称为非频繁对象集。

（4．2）利用频繁对象集生成所需要的关联规则。对于每一个频繁对象集 P，找出 P 的所有非空子集 Q，如果比率 $\varphi(P \wedge Q/S) / \sigma(P/S) \geq$最小可信度，就生成关联规则 P→Q。

下面简要介绍一下空间关联规则的算法过程：

输入：空间数据库、挖掘查询、关联约束条件集

- 数据库包括三部分：包含空间对象的空间数据库、描述非空间信息的关系数据库、概念层次集；
- 查询也包括三部分：要描述的对象类S、与任务有关的空间对象子类C_1，...C_n、有关的关系和谓语集；
- 三个约束条件：每个概念层次上的最小支持度、最小可信度、满足“临近”空间谓语的最大距离。

输出：有关对象的多种概念层次上的较强的空间关联规则

方法：空间关联发现的计算程序如下：

1）生成要描述的对象类的图层 1

2）对于相关对象的所有子类对象 Ci 同时进行如下处理：

  - 生成关联对象子类的图层 2；
  - 生成作用于对象C_i的谓词（如图层 1 与图层 2 相交，给定距离条件）；
  - 将对象C_i的谓词加入到谓词数据库中。

3）搜索谓词数据库中可信度高的空间关联规则。

空间谓词存在一个扩展的关系数据库——谓词数据库中，属性值为单值或多值集合。谓

词数据库的一个记录描述对象类S中的一个对象S_i与满足谓词关系的对象C_j的关系。

5．GIS 空间分析与空间动态建模

目前，GIS 已被广泛接受为管理、存储、查询分析有关地理空间分布的信息的成熟的工具和技术，近年来对 GIS 提出的越来越多的挑战是将这个本质上静态的系统变为支持模拟空间——时间过程的经济可行的工具，为科学研究和资源管理服务。目前大多数 GIS 软件不能支持多学科领域的动态过程计算模型，包括水文、生态等地学和生物科学，大多数 GIS 系统只是为模型提供输入参数和显示输出结果。集成 GIS 空间分析与空间动态建模困难的原因在于：

1）领域科学家、数据收集者和 GIS 技术人员存在着重要的概念差异，导致他们对现实世界抽象方式、如何用计算机组织和结构化数据的方式存在很大差异。

2） GIS 记录、存储的数据可能不是模型最适合的数据格式，有时需要进行转换。

3）除非对计算机编程非常熟悉，否则编写一个空间过程的动态模型是一个困难的很花费时间的过程。

3） 许多 GIS 软件的可视化方法并不支持对模型结果的交互式空间——时间表现。

5）满足许多科学家要求的、快速有效的模拟动态模型的开放系统模拟方法（如元胞自动机），在 GIS 实现的很少。

所谓空间动态模型是指现实世界中地球表面特定位置上的某些属性或状态因其驱动力随时间变化而发生改变的一种数学模拟。从计算的角度考虑，空间动态模型就是要计算各空间单元随其时间序列信息或驱动力的变化而产生新状态或属性值。各时刻的状态或属性可以用计算机以动画的形式进行显示，其实质仍然是图形代数和地图模型的自然延伸。但由于GIS 空间分析功能的语义限制，不能直接支持空间动态建模的操作元语，所以目前很多模型都是在 GIS 支持下，经过扩展或修改，采取与 GIS 相互结合的方式实现的。

5．1 GIS 与空间动态模型的结合方式

GIS 与空间动态模型的结合主要有三种方式：松散结合，交互界面和镶嵌结合。松散结合主要用于模型与 GIS 各自开发的情形，在这种结合中，先用 GIS 查询，预处理空间数据，然后按模型所需的格式将其输入模型，模型计算的结果再转为 GIS 文件格式，进行显示和分析。松散结合的优点在于无须改变模型的代码，不足之处在于大量的数据管理和转换工作，要求模型和 GIS 两方面的专门知识。第二种结合方式是开发交互界面。这种结合就是开发支持模型的 GIS 数据库，同时开发一个界面，既可以为模型提供输入数据，又能利用模型结果进行处理和演示，所有的数据转换是通过交互界面自动进行的。此类结合的实例最多，交互界面的优点在于提供了模型与 GIS 平台的接口，节省了大量数据输入输出工作。不足在于编程和数据管理上的投资比较大，用户修改和重写模型的难度大。第三种方式是镶嵌结合。这类结合要么是一个简单的 GIS 嵌入到一个复杂的模型系统以提供结果显示和交互控制功能，要么是用 GIS 的函数命令重写模型。这样模型和 GIS 同在一个系统中，二者共用同一个数据库，不存在数据交换问题。当模型对 GIS 的功能要求不多时常采用前一种镶嵌方法。而用 GIS 函数命令重写模型的优点在于用户可以建立自己的模型，不足在于 GIS 的命令和函数有时不能满足建立模型的所有需要。

5．2 元胞自动机简介

元胞自动机（Cellular Automata， CA）是定义在一个具有离散、有限状态的元胞组成的元胞空间上的，按照一定局部规则，在离散的时间维上演化的动力学系统。元胞自动机的基本单元是元胞（Cell），每个元胞具有一个状态，这个状态只能取有限状态集中的一个，例如“生”或“死”，或者 256 种颜色中的一种等等；这些元胞规则地排列在被称为“元胞空间”的空间格网上；它们各自的状态随着时间变化，根据一个局部的规则来进行更新，即一个元胞在某时刻的状态取决于且只取决于该元胞周围邻域元胞的状态；元胞空间内的元胞依照此局部规则进行同步的状态更新，整个元胞空间则表现为在离散的时间维上变化。
元胞自动机的最基本的组成包括元胞（Cell），元胞空间（Lattice），邻域（Neighbor），规则（Rule）。元胞自动机可以视为由一个元胞空间和定义在该空间的变换函数所组成。标准的元胞自动机是一个四元组：

$A=(d, S, N, f)$

A 代表一个元胞自动机系统；d 是一个正整数，表示元胞自动机的维数；S 是元胞的有限的离散的状态集合；N 表示一个所有邻域内元胞的组合，即包含个不同元胞状态的空间矢量，记为：

$N = (s_1, s_2, s_3, \ldots, s_n)$

n是邻域内元胞的个数；$s_i \in Z$（整数集合），$i \in (1,2,...n)$；f是变化规则，为将S^n映射到S上的一个局部转换函数。所有的元胞位于d维空间上，其位置可用一个d元的整数矢量Z^d来确定。

一般地，元胞自动机的基本模型具有五个主要特征：

1）它们由元胞的离散格局构成；

2）它们在离散时间步序内演化；

3）每一元胞的状态均在同一有限集中取值；

4）每一元胞的状态依同一确定的法则演化；

5）元胞状态的取值法则仅依赖于其自身及其周围邻域元胞的状态值。

元胞自动机虽然是产生于并行计算机结构的一种理论模型，但它还可用来描述具有很大自由度的离散系统，可视为偏微分方程离散化的理想形式。元胞自动机模型可用来模拟研究很多的现象，包括信息传递、计算、构造、生长、复制、竞争与进化等，同时，它为动力学系统理论中有关秩序、紊动、混沌、非对称、分形等系统整体行为与现象的研究提供了一个有效的模型工具。

元胞自动机模型具有很强的灵活性和开放性，它不是一个简单的数理方程，而更象是一种方法论。各领域专家可对模型的各个组成部分进行灵活的扩展，建立适合模拟各种专题现象的扩展模型，这是元胞自动机广泛应用于社会、经济、环境、地学、生物等领域的原因。尤其需要说明的是当元胞自动机在二维空间上时，元胞空间结构与栅格 GIS 数据结构高度相容，因此使用栅格 GIS 结合元胞自动机模型，可以用于离散时间和离散空间的框架下对复杂时空动态过程进行模拟。下面以一个林火模型的实例对 GIS 与元胞自动机模型的结合，模拟林火时空动态蔓延过程。

5．3 元胞自动机模拟林火蔓延模型

森林火灾是一种常见的自然灾害，当森林火灾发生时，快速准确地模拟和预测火势的蔓延和发展，对于及时部署灭火力量，搬迁居民，减少人民生命财产损失，具有重大的实际意

义。林火模型中最成功的是 Rothermel 的林火蔓延模型，模型主要考虑的因素有：

1）森林的材质：包括燃料物质载荷、燃料深度、燃料粒子密度、热容量、灭火所需湿度等。

2）燃料湿度、空气温度。

3）风速、风向。

4）地形坡度。

模型可以计算林火的蔓延速率，火灾的强度、范围和面积等。

基于该模型的框架，一些研究人员结合元胞自动机原理和 GIS，建立了新的林火模型。该模型是一个二维元胞自动机模型，基本假设条件有：森林均匀分布，材质均一；地形为平原，不考虑地形影响；风向随机，即不考虑风向影响。模型的特征如下：

元胞空间：元胞空间是将实际研究区按照一定分辨率划分的离散网格，格网单元为正方形单元，与 GIS 栅格数据结构一致。

元胞状态：每个元胞具有四种状态，0 代表无森林覆盖，1 代表未燃森林，2，代表正在燃烧森林，3 代表已烧过的森林。

邻域定义：每个元胞以 8 个相邻元胞作为其邻域元胞。

转换规则：林火蔓延规则，即在燃元胞的扩展规则。我们假定在燃元胞可以一次点燃周围邻居中的所有未燃森林，具体规则如下：

如果$n_{i,j}(t)=0$ 或 $n_{i,j}(t)=3$，则$n_{i,j}(t+1)=n_{i,j}(t)$ ；

如果$n_{i,j}(t)=2$，则$n_{i,j}(t+1)=3$ ；

如果$n_{i,j}(t)=1$ 且 $n^{'}(t)=2$（$n^{'}$为邻居元胞集合中的元素），则$n_{i,j}(t+1)=2$；

否则，$n_{i,j}(t+1)=1$。

这是一个对现实高度简化的林火蔓延模型，蔓延表现为由火点向四周成圆形扩散，基本上反映了一个理想化林火蔓延过程。实际上这个模型是一个最基本的二维通用扩散模型，可以在此基础上，加以改进、扩展，用来模拟疾病传播、污水扩散等现象。

为了更加真实的模拟林火行为过程，对上面的理想模型进行了扩展，扩展后具体模型如下。

元胞空间：与理想模型一致，只是在应用中与数字高程模型、遥感影象等数据的分辨率相匹配。

元胞状态：元胞的状态除表示森林燃烧状态外，还扩展了森林材质、湿度、地形坡度、风速风向等变量。另外为了更精确描述燃烧的过程，燃烧状态又细分成三个子状态，其中 21 表示刚刚被点燃的森林，22 表示火势旺盛的森林，23 表示由明火转向暗火的森林（逐渐熄灭）。由于不同材质的森林在各个燃烧阶段的持续时间不同，气温和湿度也会对燃烧阶段的持续时间产生影影响，所以在森林着火之前，需要确定各个森林元胞一旦着火后，处于燃烧三个子阶段的时间（可以通过 GIS 再分类命令实现）RT1，RT2，RT3。其它三个状态不变，还是 0 代表无森林覆盖，1 代表未燃森林，3 代表已烧过的森林。

邻域定义：每个元胞以 8 个相邻元胞作为其邻域元胞。

转换规则：林火蔓延规则进行了扩展。假定只有状态为 22 的元胞，燃烧正旺的森林才能点燃其邻域元胞；状态为 21 的元胞刚被点燃，火势较小不会点燃周围元胞；状态为 23 的元胞则由于火势减弱，也无法点燃周围元胞。具体的转换规则如下：

如果$n_{i,j}(t)=0$ 或 $n_{i,j}(t)=3$，则$n_{i,j}(t+1)=n_{i,j}(t)$ ；

如果$n_{i,j}(t)=21$，则有

若$RT_{i,j}>RT1_{i,j}$，则$n_{i,j}(t+1)=22$ ；

否则$n_{i,j}(t+1)=21$，$RT_{i,j}=RT_{i,j}+1$；

如果$n_{i,j}(t)=22$，则有

若$RT_{i,j}> RT1_{i,j} +RT2_{i,j}$，则$n_{i,j}(t+1) = 23$；

否则$n_{i,j}(t+1) = 22$，$RT_{i,j}= RT_{i,j}+1$；

如果$n_{i,j}(t) = 23$，则有

若$RT_{i,j}>= RT1_{i,j} +RT2_{i,j} +RT3_{i,j}$，则$n_{i,j}(t+1) = 3$；

否则$n_{i,j}(t+1) = 23$，$RT_{i,j}= RT_{i,j}+1$；

如果$n_{i,j}(t) = 1$ 且 $n'(t) = 22$（n'为邻居元胞集合中的元素），

若当前元胞被点燃的可能性超过设定的概率水平后，则$n_{i,j}(t+1) = 21$，$RT_{i,j}= 0$；

否则，$n_{i,j}(t+1) = 1$。

其中，计算森林元胞被邻域元胞点燃的可能性的计算方法是：

首先得到元胞状态为 22 的各个元胞点燃周围森林的可能性向量P_k，如图 10-15 所示：

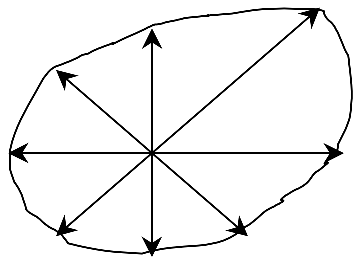


图 10-15：状态为 22 的元胞点燃周围森林的可能性向量图

然后计算当前森林元胞被点燃的可能性是它的所有相邻元胞点燃它的概率和。若超过了设定的概率水平，则可用随机数的产生方法来确定当前元胞是否被点燃。

至此，一个较为精确的林火蔓延的动态模型构建完成。这样将模型集成到 GIS 系统中，在空间数据库、遥感影象、数字高程模型等数据的支持下，提取合理的模型参数，对森林不同地点、不同天气条件下，一旦着火后的林火蔓延过程进行动态模拟和预测，对森林防火、火灾救险工作有很大的应用价值。

5．4 元胞自动机与 GIS 集成应用中的局限性

1）简单性与真实性的矛盾问题

元胞自动机是对现实世界的高度抽象和概括，它能简洁直观地模拟空间复杂系统的动态演化，但同时人们不仅会怀疑现实系统是不是这样演化的，真实性是元胞自动机模型面临的最大质疑。问题在于：元胞自动机只考虑元胞之间的局部作用，而没考虑元胞空间的宏观作用，因而忽略了现实系统的宏观作用因素；元胞自动机的因素层过于单一，元胞状态的变化取决于自身和邻域元胞的状态组合，从而忽略了其它因素的影响；标准元胞自动机的转换规则是确定性的，而现实系统的行为并非是确定性的，往往表现为某种倾向和可能性。所以在实际应用中，需要在简单性和真实性之间寻找一个平衡点，避免模型过于简单，造成结果的不真实，又要防止模型过于复杂，而失去了模型本身的意义和优势。

2）空间划分问题

元胞自动机是建立在离散、规则的空间划分基础上的，但如何确定合适的空间分辨率却是元胞自动机应用的一个难题。尤其是在多种地理实体共存的系统中，不同的实体有着不同的空间尺度，如何确定一个同一的空间分辨率，对于元胞自动机建模更为困难。另外，在不同的空间分辨率下，地理系统单元所表现出的规律也有所不同，因此，如何确定适当的分辨率，并进一步制定合理的模型规则是元胞自动机建模的一个难题。

3）时间对应问题

在元胞自动机模型中，时间是一个抽象的概念，从模型中的 T 时刻到 T+1 时刻对应的

时间单位不明确，而不同元胞在不同状态下持续的时间也很难确定。

4）转换规则定义问题

合理的规则是模型效果的关键，在元胞自动机模型中，规则是针对抽象空间划分单元的，反映了单元间局部的相互作用。此局部规则与传统的宏观规律，既有联系，又存在差别，而且找到一个确定性的规则的难度相当大。

5）与 GIS 集成的问题

GIS 系统的支持是地理系统建模研究的必要条件，虽然元胞自动机模型与栅格 GIS 在空间数据结构上存在较大的相似性，但元胞自动机是一个时空动态模型，而目前 GIS 本质上是一个静态系统，支持动态模拟的功能较弱，二者的集成有一定的困难。如何将二者紧密的集成，是阻碍元胞自动机在地理系统研究中应用的一个难题。

6．空间相互作用与位置——分配模型

6．1 地理位置

对于生产单位和服务企业来说，由于需求和供给两方面总是存在着空间上分布的差异性，因此，机构设施地理位置（Geographic Location）的选择和确定对于它们的经济效益和自身的发展具有至关重要的影响作用。

机构设施位置评价和优化，是通过对于一个设施或者一个设施网络的供给和需求两者之间的相互作用关系进行分析来实现其空间位置分布模式的优化。机构设施区位评价是对于现有服务设施的空间位置分布模式的评价，机构设施区位优化是对于其最佳位置的搜寻。克理斯塔勒的中心地理论则为供给区位优化模式的理论研究和经验方法的应用提供了基础框架，这个理论中的市场区域规模是由供货和服务的范围决定的，需求和供给两者之间的关系是以距离最小化和利润最大化为基础建立的。但是，实际应用中的中心地理论，无论是理论方法还是它的经验方法都需要根据具体问题进行具体分析。例如，与理论中假设的聚落世界相反，区域的人口集中和分散将会对该区域的范围大小产生影响；同样，消费者的消费行为的选择也会对一个服务中心的功能范围产生影响。因此，中心地理论虽然提供了一个描述和解释货物服务供给系统和需求行为两方面空间相互作用的主要框架体系，但它作为一个参考框架，还需要针对特殊情况进行扩充和调整。图 10-16 表示了中心地理论应用的概念框架。

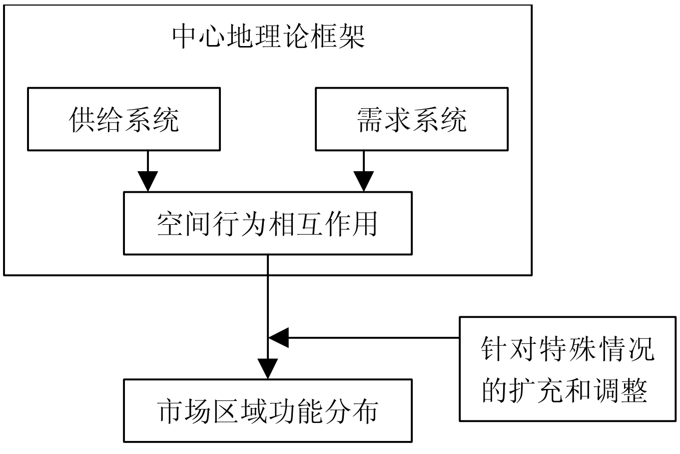


图 10-16：中心地理论应用的概念框架

中心地理论的古典模型是对包含单一目标的、且内部需求全部针对这个目标中心的具有

均质性的市场功能区域进行空间分析和模拟，而中心地理论的现代模型则要针对现实世界中那些具有不确定性的、涉及多目标消费行为和复杂供给行为的市场功能区域进行空间分析和模拟。

在对于更具现实性的市场功能行为研究基础之上，人们研究出了许多能够针对现实世界的市场功能区域进行空间分析和模拟的模型，如 Reily 的零售重力模型（Retail Gravitation）、 Batty 的裂点方程（Breakpoint Equation）、 Tobler 的价格场和作用风（Price Field and Interaction Wind）以及众多的修正式。但将重力模型及裂点方程用于分析多点中心地系统则非常艰难，由于缺乏可视性和连续性，因此，分析结果的成图也很困难。而 Tobler 方法是对中心城市场特性的一种半定量描述。下面，将重点介绍有关定位——配置问题解决方法中的一种空间线性优化模型。

6．2 空间优化模式的定义

空间优化模式用于解决位置——分配问题。位置——分配问题，是在规划重要公共设施的位置及其附属区域时产生的（公共设施，如医院、幼儿园、游戏场所、养老院、学校、警察局、消防队、急救站、管理设施等，即属于国家预算范围内的基础设施）。

一个位置一分配问题一般可表述如下：

设有一定数量的居民集中点，这些点被称为需求点（或消费点、居民点），求一定数量的供给点（某种公共设施）以及（或）供给点的需求分配，以完成某个规划目的。

如果已设需求点，求供给点，则涉及位置或定位问题（Location）。

如果已设供给点，求分配，则涉及分配或配置问题（Allocation）。

如果同时求供给点和分配，则涉及位置一分配或定位一配置问题（Location-Allocation）。

通过需求点和供给点之间的分配，供给点的附属区域也就确定了，如图 10-17 所示。

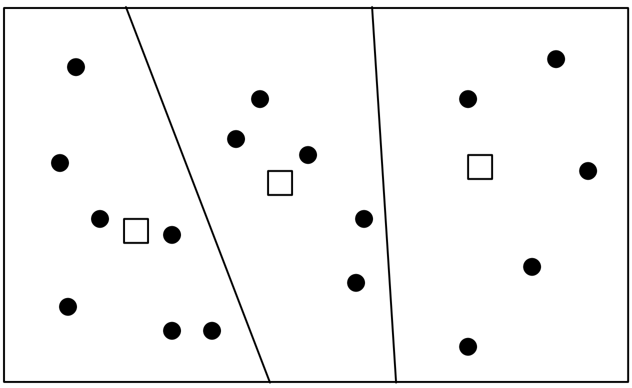


图 10-17:公共设施的位置及其附属区域

（图中实心圆表示需求点，方框表示供给点，

供应点附属区域边界为简单的直线，但实际中并非如此）

优化模式基本结构由一系列边界条件和一个（或几个，但少见）目标函数组成。在这些边界条件下，求目标函数的极大值或极小值。边界条件代表了规划目标所必须满足的划条件，它们代表了对于目标规划区域功能的基本评价；而优化目标函数（即求目标函数的极值）则代表了一个最大限度可能达到的规划目标。因此，在边界条件中体现出来的关目标函数的规划条件具有首要意义，与优化目标函数相应的规划目标的重要性则稍差一些引入目标函数的极大、极小化意义，在于得到一个定位一配置问题的明确答案（指在一定边界条件下，目标函数有数个可行答案的情况下）。

6．3 空间优化模式的分类

按照目标规划的时间范围、问题空间类型、公共设施服务方式、路途费用承担者以及公共设施使用类型等，可以将空间优化模式进行分类，如图 11-18 所示。

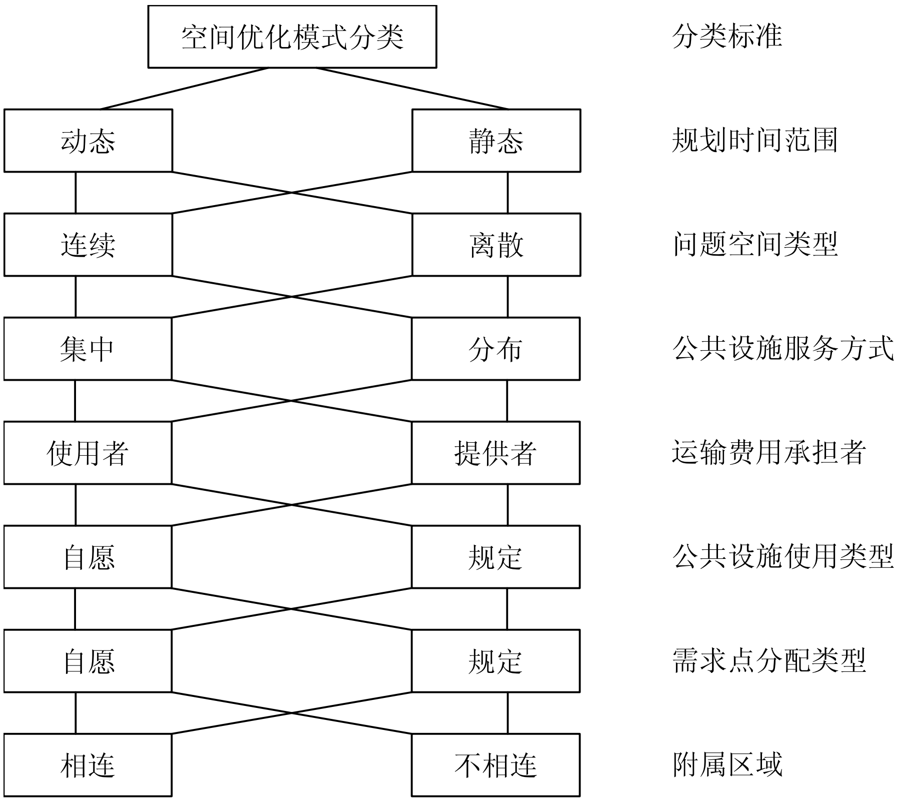


图 10-18：空间优化模式分类

6．3．1 规划时间范围

如果规划是在某个时间段（点）上解决位置一分配问题，则采用静态优化模式。如果规划时间范围包括数个时间段（点），则采用动态优化模式。动态优化模式虽然形式上易于描述，但在用于解决现实的规划问题时，由于优化目标的规模过大而难以实现，许多问题可以用静态模式加以圆满解决。

6．3．2 问题空间类型

如果所研究地区所有点都可能作为供给点，这就是一个连续性的问题。但在规划问题上，可作供给点的数目一般是有限的，因此，问题的解决可用离散模式。

规划时间范围及问题空间类型这二个维度非常重要，因为与它们相应的模式类型在形式结构的复杂性、问题解决的可能性和模式的可操作性上差别很大。下面将重点讨论“静态一离散空间优化模式”，该模式按照具体规划问题中所遇到的设施类型还可以再进行细分。

6．3．3 公共设施服务方式

如果公共设施只限于在某个确定地点为需求者提供服务，例如学校、幼儿园、养老院、医院、派出所等，需求者必须亲自前去这些公共设施接受服务，则称这些公共设施的服务方式为“集中式”。

如果公共设施所能提供的服务必须从供给点通过某种运输手段带给需求者，例如消防队、医疗急救中心、警察局等，则称这些公共设施的服务方式为“分布式”。

6．3．4 运输费用承担者

当公共设施的服务方式为“集中式”时，需求者一般直接负担自己前去这些公共设施接受服务的路途费用；当公共设施的服务方式为“分布式”时，需求者为接受服务所支付的费用一般与路途距离无关。

在“集中式”的位置——分配系统中，服务供给点的可接近程度必须从社会公正和机会均等的观点出发，而在“分布式”的位置——分配系统中工作效率问题的考虑更为重要。

6．3．5 公共设施的使用类型

一些公共设施如小学、医院、政府管理机构等，它们基于相应的法律如小学九年制义务教育、户籍管理制度等以及不可避免的客观事实如看病就医等，对于需求者具有强制性使用的性质。这种类型的公共设施由于不存在着竞争意义上的对手，因此，它们将来的需求以及设施规模都是可以预测的，称这类公共设施为规定性使用类型。

另外一些公共设施如商厦、饭店、宾馆、公园、体育馆等社会基础设施，属于需求自愿使用的设施类型。这类设施的供给与需求矛盾可能会变的非常突出，因此，它们将来的需求以及设施的规模难以预测，它们的需求者经常根据运输费用的变化而发生变化。这类设施定位时，优先考虑选择居民数量增加潜力较大的地点。

6．3．6 需求点的分配类型

和公共设施的使用类型一样，需求点的分配类型也可以是自愿或规定的。在“分布式”的位置一分配系统中，需求点的分配类型属于规定型，属于公共设施的由需求点组成的附属区域一般是确定的（既使需求者是出于自愿的）；在“集中式”的位置一分配系统中，需求点的分配类型属于自愿型，属于公共设施的由需求点组成的附属区域一般是不确定的。

比较困难的是，在需求点自愿分配的情况下估计各个供给点的附属区域及公共设施设置的规模。如何将使用者分配至供给点，可以用生产和吸引力有限的空间作用模式大概估算出来。

6．3．7 附属区域类型

需求点的规定分配与不相连的附属区域相对应，需求点的自愿分配与相连的附属区域相对应。

6．4 静态——离散空间优化模式的数学表达：线性规划

有关静态——离散的定位——配置问题是在规划重要公共设施的位置及其附属区域时出现的。静态——离散的位置——配置问题须考虑很多潜在的供给点，并且位置和附属区域的确定将在长时间内处于不变的状态。

在空间优化过程中，如果目标函数和边界条件都是线性的，则采用的数学工具是线性规划*（Linear Programming）。所谓线性规划，是指在一组线性的等式和不等式的约束下，求一个线性函数的最大值或最小值的问题。

线性规划的一般形式为：:

目标：

$\min c_1x_1+c_2x_2+\ldots\ldots+c_nx_n$

约束条件：

$a_{11}x_1+c_{12}x_2+\ldots\ldots+c_{1n}x_n\leqslant c_1$

… …

$a_{m1}x_1+c_{m2}x_2+\ldots\ldots+c_{mn}x_n\leqslant c_1$

$x_1, x_2\ldots\ldots x_n\geq 0$

下面一个例子属于产销区划，就可以用线性规划解决，有 A、B、C 三个产地，产量分别为 50、30、20，对应销地 I、II、III，需要量分别为 40、40、20。首先通过最短路径分析，得到如表 10-6 描述的运费。

表 10-6：资源分配的运费表（∞表示不连通）

	I	II	III
A	2	3	5
B	∞	1.5	8
C	2	4	∞

问题是：如何确定其供应关系，使得总运费最少。假定x_{11}表示从A到I的运量，x_{12}表示从A到II的运量，依次类推。则问题可以描述为：

Object:

$\min 2x_{11}+3x_{12}+5x_{13}+1.5x_{22}+8x_{23}+2x_{31}+4x_{32}$

Constraint:

$x_{11}+x_{12}+x_{13}=50$

$x_{22}+x_{23}=30$

$x_{31}+x_{32}=20$

$x_{11}+x_{31}=40$

$x_{12}+x_{22}+x_{32}=40$

$x_{13}+x_{23}=20$

$x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{31}, x_{32}\geq 0$

目前已经有许多成熟的解决线性规划的方案，如图解法、单纯形法等等。上述供销区划的答案如表 10-7 所示，总运费为 255。

表 10-7：资源分配的运量

	I	II	III
A	20	10	20
B	0	30	0
C	20	0	0

* 严格的说，线性规划已经不属于GIS功能的范畴，不过许多GIS软件提供此功能以实现资源分配。

基因算法（Genetic Algorithm）：

基因算法是最近提出的用于解决优化问题的方法，它通过对一个集合的变量值进行优化，得到可以定量描述优劣的目标。量化描述目标合适程度的思路来自于达尔文的进化论，在进化过程中，不同代生物染色体的基因的内容和形式都在发生变化，使得生物更能够适合于生存环境。在基因算法中，是通过不同方式的组合以及修改数值以改变变量，这类似于生物的配对和变异。通过比较合适指标，测试改变变量后的计算结果，保留“有利的”组合，并进行进一步的修改。一种理想的结果是：保留下来的改变优于其它的组合和改变，通过不断的迭代过程，可以使得一个优化的结果。基因算法的成败依赖于以何种方式来模拟各种形式的染色体，并使其不断进化，以增加它在最优解中的存活几率。存活下来的染色体一般都对应于一个局部极值，而后者可以导致全局的最优解。

在一个基因算法中，要确定的内容包括：

1）基因操作，如配对和变异；

2）策略性的参数：如配对和变异发生的几率；

3）表现模式：如对基因的编码等等。

在 GIS 中，基因算法可以应用于空间相互作用（Spatial Interaction）建模，后者可以应用于公共设施，如商店、医院等的区位。

上北下南，绿色代表平原，棕色代表高山，蓝色代表海洋。

布拉格是粉红色的，佛罗伦萨是蓝色的。

地图是一个令人想入非非的东西，它的容量如此之大，足够存储我们过于丰富的情感。

佚名

第十一章 空间数据表现与地图制图

导读：空间数据的表现功能贯穿于 GIS 空间数据处理的始终，功能的许多方面，包括符号、专题信息表达、图幅配置、制图综合等，都来源于地图制图，这也表明了地理信息系统和制图学的渊源关系。从某种角度上来讲，地理信息系统可以称为“动态的地图”，它提供了比普通地图更为丰富和灵活的空间数据表现方式，如动态信息表达、虚拟现实等等。但是，从另外的角度来看，制图学又是一门艺术，单纯依靠计算机系统是难以制作完美的地图的，它需要人工的交互干预，如制图综合，目前还没有任何 GIS 软件能够很好地解决这些问题。

本章最后介绍了可视化的内容，可视化与信息的感知和传输有着密切的关系，地学可视化的研究、以及在 GIS 中应用可视化技术，对于空间数据的广泛应用有着重要的意义。

1. 概述——地理信息系统数据表现与地图学

地图是记录地理信息的一种图形语言形式，从历史发展的角度来看，地理信息系统脱胎于地图，并成为地图信息的又一种新的载体形式，它具有存储、分析、显示和传输的功能，尤其是计算机制图为地图特征的数字表达、操作和显示提供了成套方法，为地理信息系统的图形输出设计提供了技术支持；同时，地图仍是目前地理信息系统的重要数据来源之一。但二者又有本质之区别：地图强调的是数据分析、符号化与显示，而地理信息系统更注重于信息分析。地图学理论与方法对地理信息系统的发展有着重要的影响，特别是地理信息系统中的空间数据表现许多技术来源于地图学，但是，其表现手段更加灵活和丰富。

计算机制图是地图学中较新的一个领域，也是地理信息系统输出、显示结果的一个重要部分。计算机制图的主要内容是将地图设计和绘制通过计算机程序和相应的绘图硬件来实现。地图符号已经可以完全由计算机绘制。对地图特征的许多处理，如分类、空间插值等均可用计算机来实现。利用计算机对线状要素进行制图综合也已经取得很大进步。但是，利用计算机进行制图综合的结果仍比较机械，需要将更多的专家经验计算机化，特别是用计算机同时对多种内容进行地图综合还不太成熟。计算机图形学是利用计算机处理图形信息以及借助图形信息进行人——机通讯处理的技术，是 GIS 算法设计的基础，GIS 是随着计算机图形学技术的发展而不断发展完善的。

1．1 构成地图学与地理信息系统的数学法则

地图对它所表现的地面上的景物，应能按比例尺衡量。比例尺是地图上直线长度与地面上相应距离的水平投影长度之比。如地图上注有比例尺 1：50000，表示图上地物的长度相当于相应地面地物长度的 1/50000。这个尺度对大比例尺地图来说，基本上适合图上各个部

分。因此，可以认为，比例尺是地图线形缩小程度的标志，它是构成地图数学要素的基本组成部分之一。地图上各种地物之间的关系，要求按数学法则构成，这就是先将地球自然表面的景物垂直投影到地球椭球面（或球面）上，再将地球椭球面（或球面）按数学法则投影到平面上而构成地图。这种按数学法则将地球椭球面（或球面）转绘到平面上的方法，叫地图投影。按这种方法建立的数学基础，才能使地球表面上各点和地图平面上的相应各点保持一定的函数关系，从而才能在地图上准确地表达空间各要素的关系和分布规律，才可能反映出它们之间的方向、距离和面积，使地图具有区域性和可量测性。

1．2 地图学和地理信息系统的符号

地图符号是表达地图内容的基本手段，它不仅能表示事物的空间位置、形状、质量和数量特征，而且还可以表示各事物之间的相互联系及区域总体特征。图形符号是由形状、尺寸和颜色三个基本因素所组成，具有系统化的特点；注记是地图符号的一个重要部分，它也有形状、尺寸和颜色的区别。根据事物分布的特点，地图符号有面状、点状和线状之分。这些符号运用在表示专题要素的分布、质量和数量特征以及动态变化时，形成了许多表示法：范围法、质底法、量底法、等值线法、定点符号法、线状符号法、动线法、点值法、定域统计图法、色级统计图法以及定位统计图法。表示地形的主要方法有等高线法、分层设色法、晕渲法等。

地图符号是表示地图内容的基本手段，它由形状不同、大小不一、色彩有别的图形和文字组成。地图符号是地图的语言，是一种图形语言。它与文字语言相比较，最大的特点是形象直观。就单个符号而言，它可以表示某个事物的空间位置、大小、质量和数量特征；就同类符号而言，可以反映各类要素的分布特点；而各类符号的总和，则可以表明各类要素之间的相互关系及区域总体特征。因此，地图符号不仅具有确定客观事物空间位置、分布特点以及质量和数量特征的基本功能，而且还具有相互联系和共同表达地理环境各要素总体的特殊功能。

地图符号的形成过程，可以说是一种约定的过程，经过很长时间的检验，有约定而达到俗成的程度，为广大用图者所熟悉和承认。地图符号的作用，在于它能保证所表示的客观事物空间位置具有较高的几何精度，从而提供了可测量性；能用不依据比例符号或半依据比例符号表示出事物的质量和数量特征。

地图是运用易被人们感受的图形符号表示地面景物的，使用符号具有以下功效：

1）有选择地表示地理环境中的主要事物，因而在较小比例尺的地图上所表现的地面情况，仍能一目了然，重点突出。对于那些由于缩小而不能按比例尺表示的重要地面景物，可用不依比例的符号夸大表示；

2）用平面的图形符号表示地面的起伏状况，也可以说是在二维平面上，能够表达出三维空间状况，而且可以量测其长度、高度和坡度等；

3）除了用符号表示出地面景物的外形，还能表示出景物的看不见的本质特征，如在海图上可以表示出海底地形、海底地质、海水的温度和含盐度等；

4）用符号可以表示出地面没有外形的许多自然和社会经济现象，如气压、雨量、政区和人口移动等。此外还可以表现出事物间的联系和制约关系，如森林分布和木材加工工业之间的联系。地图上还有起说明作用的文字和数字，它们也是地图的重要组成部分，用以标明地面景物的名称、质量和数量。

1．3 地图学和地理信息系统的制图综合

制图综合是对制图区域客观事物的取舍和简化。经过概括后的地图可以显示出主要的事物和本质的特征。地图的比例尺、用途和主题，制图区域的地理特征以及符号的图形尺寸是影响地图概括的主要因素，地图概括主要表现在内容的取舍、数量简化、质量化简和形状化简等方面。

地图最重要、最基本的特征是以缩小的形式表达地面事物的空间结构，这个特征表明，地图不可能把地面全部事物毫无遗漏地表示出来。地图上所表示的地面状况是经过概括后的结果。地图和实际地面相比，是缩小的。地图上所表现的地面景物，从数量上看是少的，从图形上看是小了、简化了。这是因为地图上所表现的内容都是经过取舍和化简的。由 1：10000 比例尺缩小到 1：50000 比例尺的地图，对原来的内容如不进行取舍和化简，缩小后的地图既不清楚又不易读。这种把实地景物缩小或把原来较详细的地图缩成更小比例尺地图时，根据地图用途或主题的需要，对实况或原图内容进行取舍和化简，以便在有限的图面上表达出制图区域的基本特征和地理要素的主要特点的理论与方法，称为地图综合（地图概括）。制图综合就是对客观事物进行取舍和化简，取舍就是从大量的客观事物中选出最重要的事物表示在图上，而舍去次要的事物；化简就是对客观事物的形状、数量和质量特征的化简；形状化简是去掉轮廓形状的碎部，以突出事物的总体特征；数量和质量特征的化简就是减少分类和分级的数量，以缩小与客观事物的差别。当然取舍和化简不是任意的，而是根据地图的比例尺、用途和制图区域的地理特征，对地图上各要素及其内在联系加以分析研究，选取的目的是强调主要的事物和本质的特征，而舍去次要的事物和非本质的特征。制图综合可分为比例概括和目的概括。前者是由于地图比例尺缩小，图形也就缩小，有些图形缩小到难以清楚地表达出来，从而必须选取和化简；后者是因为客观事物的重要性并不完全决定于它的图形大小，故而它的选取和化简也不完全由比例尺决定，还要根据编图者对客观事物重要性的判断来确定是否选取和化简。

从地理信息系统的发展过程可以看出，地理信息系统的产生、发展与制图信息系统存在着密切的联系，两者的相通之处是基于空间数据库的表达、显示和处理。从系统构成与功能上看，一个地理信息系统具有机助制图系统的所有组成和功能，并且地理信息系统还有数据处理的功能。地图是一种图解图像，是根据地理思想对现实世界进行科学抽象和符号表示的一种地理模型，是地理思维的产物，也是实体世界地理信息的高效载体，地图可以从不同方面、不同专题，系统地记录和传输实体世界历史的、现在的和规划预测的地理景观信息。

2．地图的符号

2．1 地图符号的实质

2．1．1 概念

地图符号（Symbol）是地图的语言，它是表达地图内容的基本手段。地图符号是由形状不同、大小不一和色彩有别的图形和文字组成，注记是地图符号的一个重要部分，它也有形状、尺寸和颜色之区别。就单个符号而言，它可以表示事物的空间位置、大小、质量和数

量特征；就同类符号而言，可以反映各类要素的分布特点；而各类符号的总和，则可以表明各要素之间的相互关系及区域总体特征。

地图符号可以指出目标种类（如公路）及其数量特征和质量特征（如公路行车部分的铺面种类和宽度），并且可以确定对象的空间位置和现象的分布（如人口密度等）。

2．1．2 分类

1）按照符号的定位情况分类

可以将符号分为定位符号和说明符号。

定位符号是指图上有确定位置，一般不能任意移动的符号，如河流、居民地及边界等，地图上的符号大部分都属于这一类；

说明符号是指为了说明事物的质量和数量特征而附加的一类符号，它通常是依附于定位符号而存在的，如说明森林树种的符号等。它们在图上配置于地类界范围内，但都没有定位意义。

2）按照符号所代表的客观事物分布状况分类

可以把符号分为面状符号、点状符号和线状符号。

面状符号是一种能按地图比例尺表示出事物分布范围的符号。面状符号是用轮廓线（实线、虚线或点线）表示事物的分布范围，其形状与事物的平面图形相似，轮廓线内加绘颜色或说明符号以表示它的性质和数量，并可以从图上量测其长度、宽度和面积，一般又把这种符号称为依比例符号（图 11-1）。

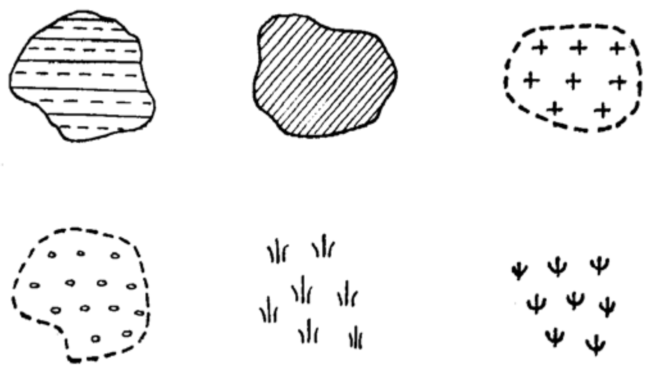


图 11-1：面状符号

点状符号是一种表达不能依比例尺表示的小面积事物（如油库等）和点状（如控制点）所采用的符号。点状符号的形状和颜色表示事物的性质，点状符号的大小通常反映事物的等级或数量特征，但是符号的大小与形状与地图比例尺无关，它只具有定位意义，一般又称这种符号为不依比例尺符号（图 11-2）。

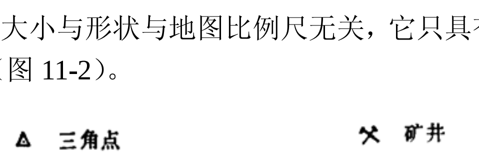


图 11-2：点符号

线状符号是一种表达呈线状或带状延伸分布事物的符号，如河流，其长度能按比例尺表示，而宽度一般不能按比例尺表示，需要进行适当的夸大。因而，线状符号的形状和颜色表

示事物的质量特征，其宽度往往反映事物的等级或数值。这类符号能表示事物的分布位置、延伸形态和长度，但不能表示其宽度，一般又称为半依比例符号（图 11-3）。

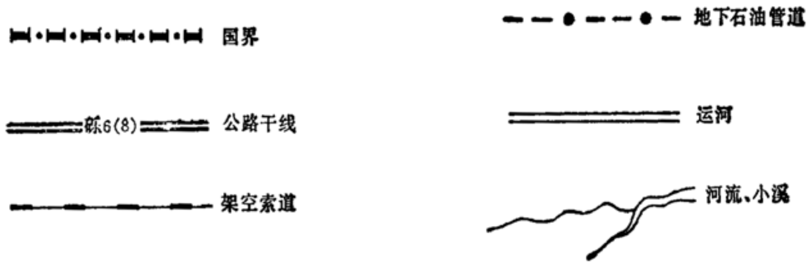


图 11-3：线状符号

2．2 地图符号的构成特点

2．2．1 符号的构成要素

地图上的形状、尺寸和颜色是构成符号的三个基本要素。

1）符号的形状

符号的形状主要是表示事物的外形和特征，面状符号的形状是由它所表示的事物平面图形决定的，点状符号的形状往往与事物外部特征相联系，线状符号的形状是各种形式的线划，如单线、双线。

2）符号尺寸

符号尺寸大小和地图内容、用途、比例尺、目视分辨能力、绘图与印刷能力等都有关系，不同的比例尺的地图，其符号大小也有所不同。

3）符号的颜色

符号的颜色可以增强地图各要素分类、分级的概念，简化符号的形状差别，减少符号数量，提高地图的表现力。使用颜色主要用以反映事物的质量特征、数量特征和等级。

2．2．2 地图符号的系统化

由形状、尺寸和颜色变化组成的各种地图符号并不是孤立的，它们具有内在的联系。通过符号的变化可以把地图内容的分类、分级、重要、次要等不同情况表达出来。

2．2．3 地图符号的感受效果

地图是由在不同位置上的符号图形所组成的，符号的复杂排列能引起视觉的不同感受。读图时可获得整体感、差异感、立体感等不同的感受效果。整体感主要依靠符号之间不存在明显的差异而形成的；差异感是由不同组成要素在图形间产生明显区别而形成的；立体感主要是利用透视的方法，通过改变符号的尺寸、颜色的亮度等来实现的，如晕渲法就是利用光影，模拟地表面受光的强弱，用浓淡不同的色调显示地形起伏而产生立体感。

地图符号使用了不同的视觉变量，如尺寸、形状、灰度、纹理、方向和颜色等，实现了不同的感受效果。（图 11-4）

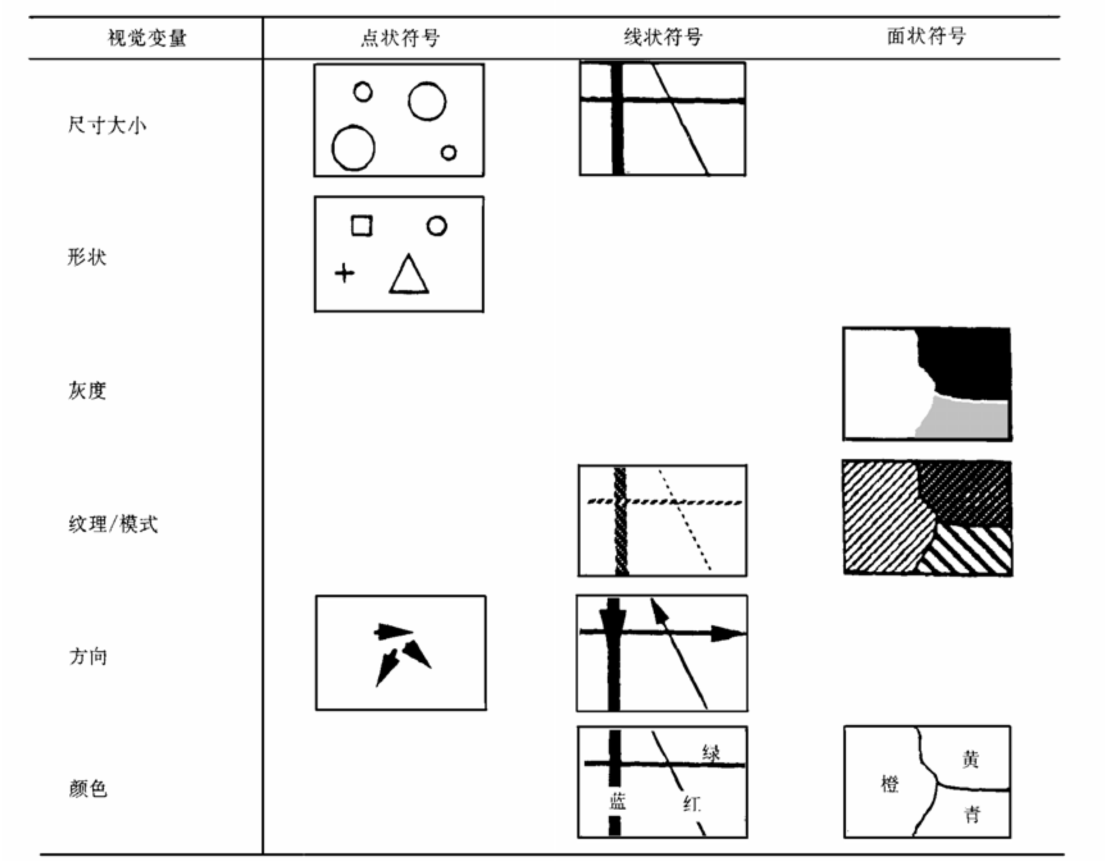


图 11-4：地图符号效果

2．3 地图上的注记

2．3．1 地图注记的种类

地图上的文字和数字总称为地图注记，它是地图内容的重要部分。注记并不是自然界中的一种要素，但它们与地图上表示的要素有关，没有注记的地图只能表达事物的空间概念，而不能表示事物的名称和某些质量和数量特征。地图上的注记可分为名称注记、说明注记和数字注记三种。

1）名称注记 说明各种事物的专有名称，如居民点名称；

2）说明注记 用来说明各种事物的种类、性质或特征，用于补充图形符号的不足，它常用简注表示；

3）数字注记 用来说明某些事物的数量特征，如高程等。

2．3．2 注记的字体、字级和颜色

地图上常用不同的字体表示不同的事物，常用的字体有宋体、等线体、仿宋体和横线体；

地图上注记尺寸的大小，以照相排字机注明的规格为标准，在一幅图上，按照事物的重要程度和意义，采用不同的字级，以便使注记大小与图形符号相对应；注记的颜色只有色相的变化，颜色的选用要与注记所表示的事物类别相联系。

2．3．3 注记的排列和配置

地图上注记数量较多，它们可以位于地图中的任一部分，但是注记的排列和配置是否恰当，常常会影响读图的效果。汉字注记通常有水平字列、垂直字列、雁形字列（注记的字向指向北方或图廓上方）和屈曲字列（注记的字向与注记文字中心线垂直或平行）等等。注记配置的基本原则是不应该使注记压盖图上的重要部分。注记应与其所说明的事物关系明确。对于点状地物，应以点状符号为中心，在其上下左右四个方向中的任一适当位置配置注记，注记呈水平方向排列；对于线状事物，注记沿线状符号延伸方向从左向右或从上向下排列，字的间隔均匀一致，特别长的线状地物，名称注记可重复出现；对于面状事物，注记一般置于面状符号之内，沿面状符号最大延伸方向配置，字的间隔均匀一致。

图 11-5 表示了点、线、面注记的不同配置方式。

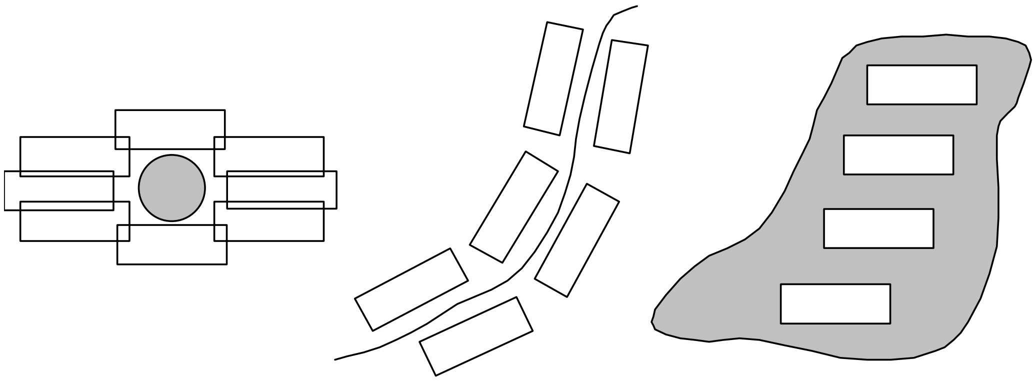


图 11-5：几种可能的注记配置方式

为了确定地物注记的位置，要进行空间关系的判断，因此 GIS 可以为地图制图提供自动标注功能，一个实现自动注记放置的系统要具有以下功能：

1）确定地图上的要素以及相应的注记文字；

2）对空间数据进行搜索和实现；

3）产生试验性的注记点；

4）选择较好的注记位置。

由于注记只是对地物的描述，因此在地图上，注记不能遮盖地物，注记之间也不能相互重叠，所以，也可以利用 GIS 进行注记是否重叠的判断。在进行注记时，由于图面荷载的原因，不可能对所有的地物进行标注，这需要进行选择，通常的选择方法是选择相对重要的地物，这也可以通过对地物属性的排序来实现。

3．专题信息表现

在大多数地理信息系统中，管理的空间数据是矢量格式的地物对象，这些地物对象不仅具有空间位置特征，而且具有非空间的属性数据。在表现这些地物对象时，除了显示空间位置以外，同时还可以以特定的方式显示某个或多个相关的属性，生成专题地图。

专题地图除了采用普通地图某些表示方法并使其得到进一步发展外，本身还需要有专门反映各种要素性质、数量、空间分布和时间变化的表示方法，以便使读者明确对专题内容要

素的科学分析。

在专题地图中，各种制图对象的基本形状是由点、线、面及其过渡形态组成的，并以此反映现象的分布特点、现象的变化时刻、质量和数量的特征及综合特征。因此，在选择表示方法时，既可以根据专题制图对象的分布方式进行选择（表 11-1），也可以按它们的分布特点进行选择（表 11-2）。

表 11-1：按照专题制图对象的分布方式选择符号

分布方式	表示分布范围	表示类别	表示数量	表示动态
点	定位符号	符号形式： 几何符号 文字符号 象形符号 符号颜色	符号大小（用深浅不同的颜色表示主次） 数字注记	定点扩展法 定位图表法
线	线状符号	线的形式 线颜色	线条粗细，长短，附加不同数目的小短线	动线法
面	面状符号（范围法）	质底法	点值线法，分级统计图法，图形统计图法，等值线法和分层设色法	表示扩展的范围法，扩展图形的图形统计图法

表 11-2：按照专题制图对象的分布特点选择符号

分布特征	状况	表示方法
固定一点	一定瞬间的状态移动	符号法、运动线法
	空间变化	符号法、定位图表法
固定一线	一定瞬间的状态移动	线状符号法、线状符号和运动符号结合
	空间变化	线状符号法重叠
固定面积	一定瞬间的状态移动	质底法、等值线法、定位图表总和、范围法、区划法结合等值线法、运动符号
	空间变化	等值线法、定位图表和符号总和
散布	一定瞬间的状态移动	点子法、质底法、区划法、图表统计法和色彩统计法、区划搭配运动符号
	空间变化	点子法和区划法结合、图表统计法和色彩统计法
连续分布	一定瞬间的状态移动	质底法、等值线法、定位图表总和、地形塑拟表示运动符号
	空间变化	等值线法、定位图表总和

3．1 专题地图的分类和内容

3．1．1 专题地图含义

专题地图是突出地表示一种或几种自然现象和社会经济现象的地图。按内容可分为三大类：自然地图、社会经济地图和其他专题地图。

自然地图表示自然界各种现象的特征、地理分布及其相互关系，如地质图、水文图等；社会经济地图表示各种社会经济现象的特征、地理分布及其相互关系，如人口图、行政区划图等；其他专题地图，指不属于上述二类的专题地图，如航海图、航空图等。

3．1．2 专题地图的内容

专题地图的种类很多，但大都是由地理基础和专题内容组成的。地理基础即普通地图上的一部分内容要素，如经纬网、水系、居民点、交通线、地势等。地理基础作为编绘专题内容的骨架，并表示专题内容的地理位置和说明专题内容与地理环境的关系。专题地图上表示哪些地理基础要素和详细程度如何，是根据专题内容的不同而有所不同。

专题内容，从资料来讲，一是将普通地图内容中一种或几种要素显示得比较完备和详细，而将其他要素放到次要地位或省略，如交通图等；二是包括在普通地图上没有的和地面上看不见的或不能直接测量的专题要素，如人口密度图。

3．2 面状专题内容的表示方法

地面上真正的点状事物很少，一般都占有一定的面积，只是大小不同。点状分布要素指那些占据的面积较小，不能按比例尺表示，又要定位的事物。对于点状分布要素的质量特征和数量特征，可以用点状符号表示。

在地面上呈线状或带状分布的事物很多，如交通线、河流及边界线等，对于这些事物的分布质量特征和数量特征可以用线状符号表示。

面状专题内容的表示方法，最常用的有：等值线法、质底法、范围法、点值法、符号法、动线法、统计图法等。

3．2．1 等值线法

等值线系指在地图上通过表示一种现象的数量指标的一些等值点的曲线，如等高线、等温线。等值线法宜用于表示地面上连续分布而逐渐变化的现象，并说明这种现象在地图上任一点的数值或强度。等值线的数值间隔原则上最好是一个常数，以便判断现象变化的急剧或和缓，但也有例外。等值线间隔的大小首先决定于现象的数值变化范围，变化范围越大（以等高线为例，地貌高程变化越大），间隔也越大，反之亦然。如果根据等值线分层设色，颜色应由浅色逐渐加深，或由寒色逐渐过渡到暖色，这样可以提高地图的表现力（图 11-6）。

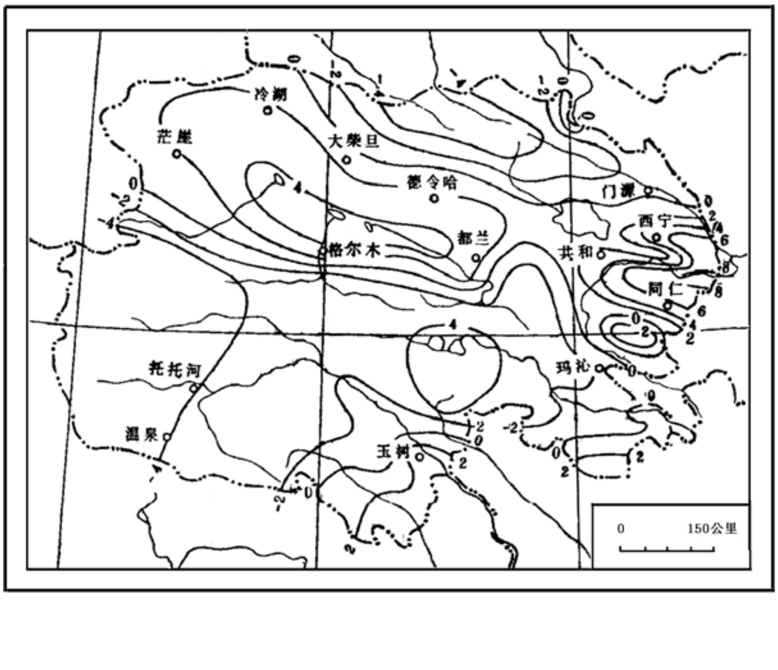


图 11-6：等值线

3．2．2 质底法

质底法又名底色法，用于将区域划分为质量相同的地段。见图 11-7。

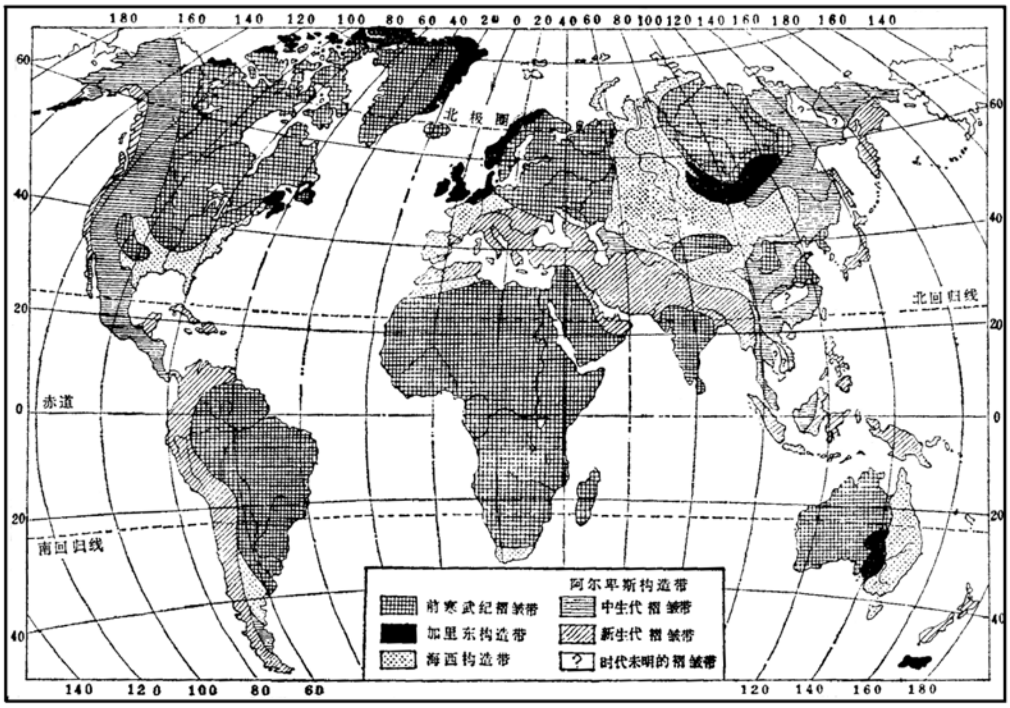


图 11-7：质底法示例

由于质底法广泛应用各种颜色，所以有时称之为底色法。首先按现象的性质进行分类或分区，制成图例，在地图上绘出各分类界线，然后把同类现象或属于同一区划的现象绘成同一颜色或同一的晕纹。这种方法可以用于表示地表面上的连续面状现象（如气象现象）、大面积分布的现象（如土壤覆盖）或大量分布的现象（如人口）。质底法的优点是鲜明美观，缺点是不易表示各类现象的逐渐过渡，而且当分类很多时，图例比较复杂，必须详细阅读图例时才能读图。注意质底的两种颜色系统不应该相互重叠，但是底色可以与晕线重合。另外，质底法易于与其他表示方法结合使用。

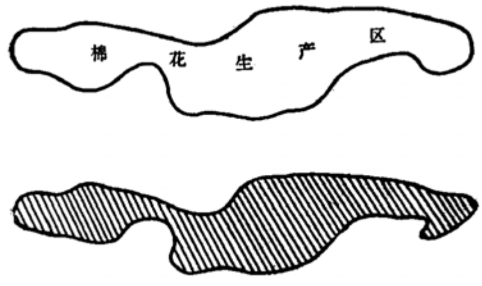
3．2．3 范围法

范围法又名区域法，用于表示某种现象在一定范围内的分布。范围法分为精确范围法（见图 11-8-1）和概略范围法（见图 11-8-2），前者有明确的界线，可以在界线内着色或填绘晕纹或文字注记；后者可用虚线、点线表示轮廓界线，或不绘轮廓界线，只以文字或单个符号表示现象分布的概略范围。

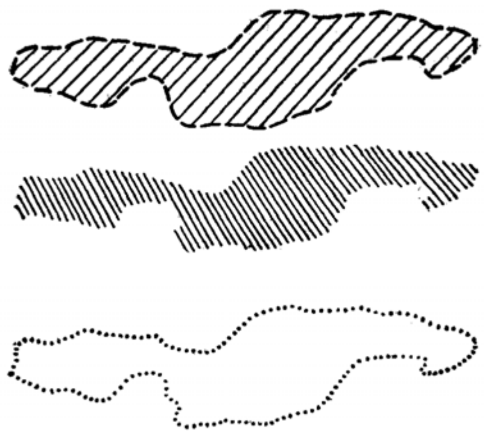
在地图上表示范围可以采用各种不同的方法：用一定图形的实线或虚线表示区域的范围；用不同颜色普染区域；在不同区域范围内绘以不同晕线；在区域范围内均匀配置晕线符号，有时不绘出境界线；在区域范围内加注说明注记或采用填充符号。

范围法与质底法的区别，在于所表现的现象不布满整个编图区域；不一定有精确的范围

界线。



(a)



(b)

图 11-8:（a）精确范围法，（b）概略范围法

3．2．4 点值法

点值法是在图上用小点表示现象的分布和数量，它适用于表示分布不均匀的现象。从图上点的疏密就可看出现象的集中或分散的程度。布点时要确定点的大小及其代表的值，在最稠密的地方，点可以近于紧接但不能重迭，在最稀疏的地方，也有点的表示；在其他地方，则依比例显示出点的疏密。（图 11-9）



图 11-9：点值法示例

点值法是质底法和范围法的进一步发展。质底法和范围法只能反映现象的分布范围及其质量特征，点值法则可以表明现象的分布和数量特征。点值法有两种方法：一是均匀布点法，即在一定的区划单位内均匀地布点；另一是定位布点法，即按照现象实际所在地布点。

点值法的优点是简单明了，如果恰当地采用不同颜色和不同形状的点，既可以表示现象的数量特征，也可以表示它的质量特征和发展概况。

3．2．5 符号法

符号法就是用各种不同形状、大小和颜色的符号表示现象的分布及其数量特征和质量特征。通常以符号的大小表示数量的差别，形状和颜色表示质量的差别，而将符号绘在现象所在的位置上。符号法可分为个体符号法和和线状符号法。

1）个体符号法

是地图表示法中的一种特殊方法，用于表示不依地图比例尺或所占面积小于地图符号本身的一些地物的位置，并且通常用于表示按点定位的现象。单个符号按其形状可分为几何符号、文字符号和象形符号。由于文字符号和象形符号不能精确地表示位置以及数量的差异，所以地图上应用最广的是几何符号。几何符号具有简单的集合图形，这种符号便于绘制，能准确地指明地物位置，便于比较大小；文字符号用表示现象名称的头一、二个字母代表。文字符号的应用受到一定限制，因为采用字母使地图不易阅读，不能指示地物的准确位置，不能按符号大小对比；象形符号相似于所表示对象的图形，可以进一步分为象征符号和实体符号。前者的形状与所表示地物有某种程度的联系；后者用实体图形表示相应的事物分布。

2）线状符号法

线状符号既可用于表示几何概念的线划，如分水岭、地面上确定的各种境界线（如国界线），又可以用于表示线状分布的不能依地图比例尺表示其宽度的地物，如河流与道路等。有时还可以用线状符号强调图上按面积表示的地物的主要方向，如山脊线、山脉走向线等。（图 11-10）

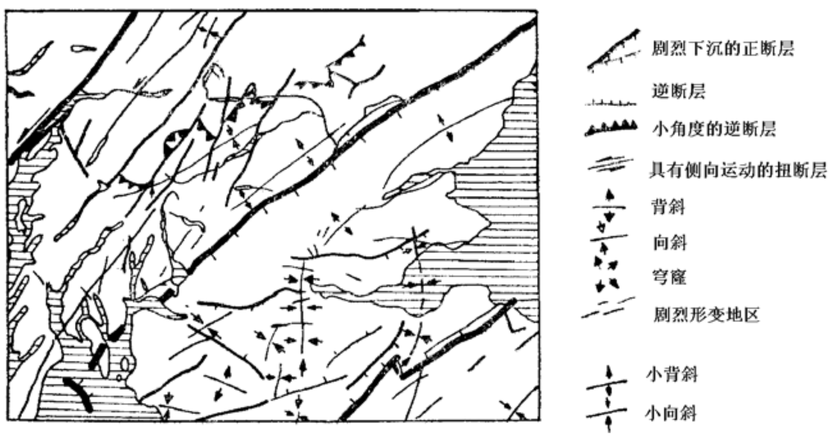


图 11-10：线状符号法示例（地质构造线）

3．2．6 动线法

动线法一般是用箭形符号表示现象的运动路线和方向，如人口迁移路线、洋流和货运路线等。箭头和箭体上部的方向应保持一致，箭头的两翼应保持对称。箭形的粗细或宽度表示洋流的速度强度或货运的数量；箭形的长短表示风向、洋流的稳定性；首尾衔接的箭形表示运动的路线，动线法表示面状事物的移动。

3．2．7 统计图法

统计图法一般根据编图区域内各区划单位或典型地点的统计资料，用地图形式表达出来。可以分为图形统计图法、分级统计图法和定位统计图法三种。

1）图形统计图法

图形统计图法是根据各区划单位的统计资料作成图形或图表，绘在地图上该区划单位之内。这种方法与符号法的区别是，它反映一个区划范围的现象，而不是一个点上的现象。图形统计图宜于表示绝对的数量指标。

在地图制图中采用较多的是：线状统计图形——柱状和带状等，其长度与所比较的数值成正比；面积统计图形——正方形、圆圈等，其面积大小与所比较的数值成正比；立体统计图形——立方体、圆球等，其体积与所比较的数值成正比（图 11-11）。

2）分级统计图法

分级统计图法表示一定区域单位范围内某种制图现象平均密度的方法。分级统计图法是按照各区划单位的统计资料，根据现象的密度、强度或发展水平划分等级，然后依据级别高低，在地图上按区划分别填绘深浅不同的颜色或疏密不同的晕线，以显示各区划单位间的差异。分级时可采用等差的、等比的、逐渐增大的或任意的。分级统计图适于表示相对的数量指标。该种统计图法的优点在于绘制简单，阅读容易。

3）定位统计图法

定位统计图法是将固定地点的统计资料，用图表形式绘在地图上的相应地点，以表示该地某种现象的变化。常用的图表有柱状图表、曲线图表、玫瑰图表等。

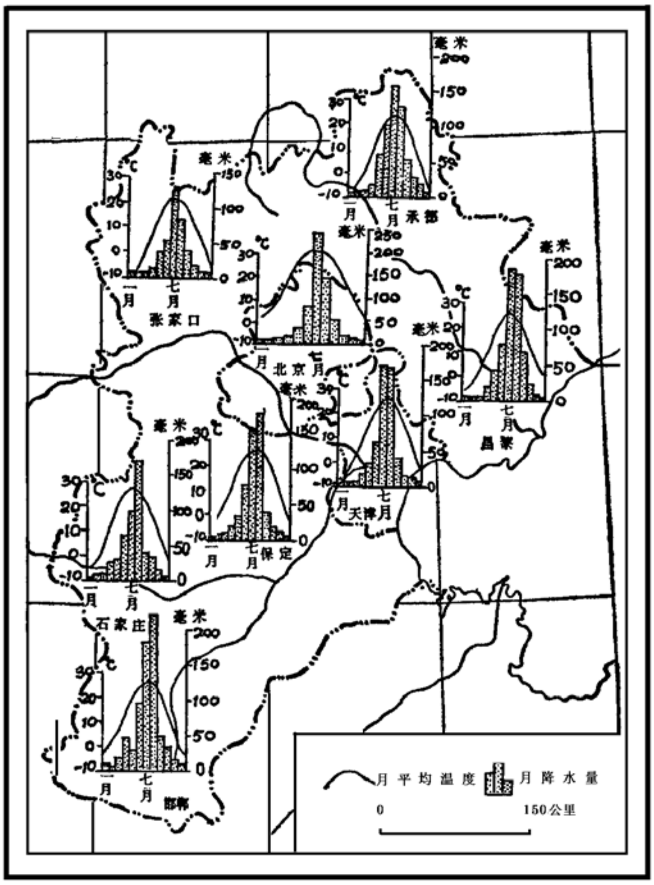


图 11-11：统计图法示例

3．2．8 DEM 表示法

数字高程是一组关于地表面位置布局的高程量测数据，主要应用于某个区域的地形分析，前面的几种方法只能反映现象的水平分布特征，对垂直分布特征是无能为力的。数字高程表示法广泛地用于工程、规划和军事领域。

1）地貌晕渲图（Shaded Relief Map）及其与专题地图叠置

为了增加丘陵和山地地区描述高差起伏的视觉效果，制图工作者成功地运用了一种"阴影立体法"即地貌晕渲法。用这种技术绘制的图件看起来很动人，但费用太高，晕渲法的质量和精度很大程度上取决于制图工作者的主观意识和技巧。

数字地形图投入生产并加以应用后，地貌晕渲便能自动、精确地实现。自动晕渲的原理是基于"地面在人们眼里是什么样子，用何种理想的材料来制作，以什么方向为光源照明方向"等的考虑。制图输出时如果用灰度级和连续色调技术表示明暗程度，得到的成果看起来与航片十分相似。实际上，从高程矩阵中自动生成的地貌晕渲图与航片有许多不同之处。主要表现在：

第一，晕渲图不包括任何地面覆盖信息，仅仅是数字化的地表起伏显示；

第二，光源一般确定为西北 45 度方向，航片的阴影主要随太阳高度角变化；

第三，晕渲图通常都经过了平滑和综合处理，因而没有航片上显示出的丰富的地形细节。

自动地貌晕渲图的计算非常简单，首先是根据 DEM 数据计算坡度和坡向。然后将坡向数据与光源方向比较，面向光源的斜坡得到浅色调灰度值，反方向的得到深色调灰度值，两者之间得到中间灰度值。灰度值的大小则按坡度进一步确定。

计算晕渲图的主要研究集中于坡面反射率的定量描述，由于计算反射率的公式都较复杂，因此，将坡度和坡向转换成反射量，常用建立查找表的方法来解决，使计算和处理更为有效（图 11-12）。

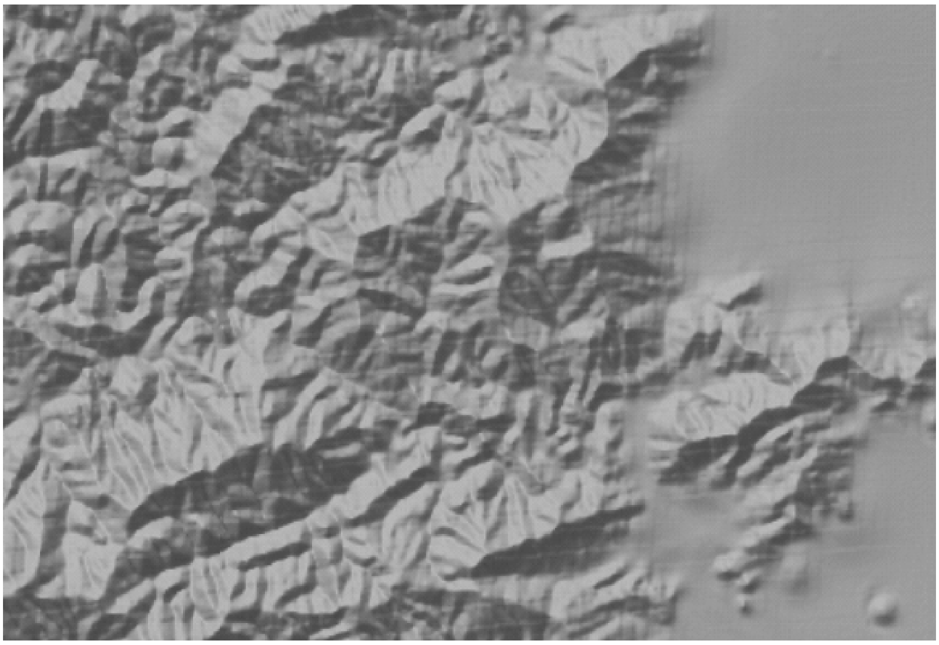


图 11-12：地貌晕渲图示意

晕渲图本身可以描述地表三维状况，而且在地形定量分析中的应用不断扩大。如果把其它专题信息与晕渲图叠置组合在一起，将大幅度提高地图的使用价值。例如，运输线路规划图与晕渲图叠加后大大增强了直观感等，这是传统方法不能实现的。

2）三维曲面

许多 GIS 软件都提供了利用 DEM 数据生成三维投影的功能，通常三维透视图有两种形式，即平滑的三维曲面和网状图（Fishnet Diagram）（图 11-13，图 11-14），其中在三维曲面图上还可以叠加其它信息，最典型是叠加同一区域的遥感图像或者该 DEM 生成的晕渲图，以增强其“立体感”。生成三维透视图，需要设置视点位置、视角等参数，因此可以通过设置不同的参数，生成连续的三维透视图，并组成动画，达到“虚拟飞行”的效果。

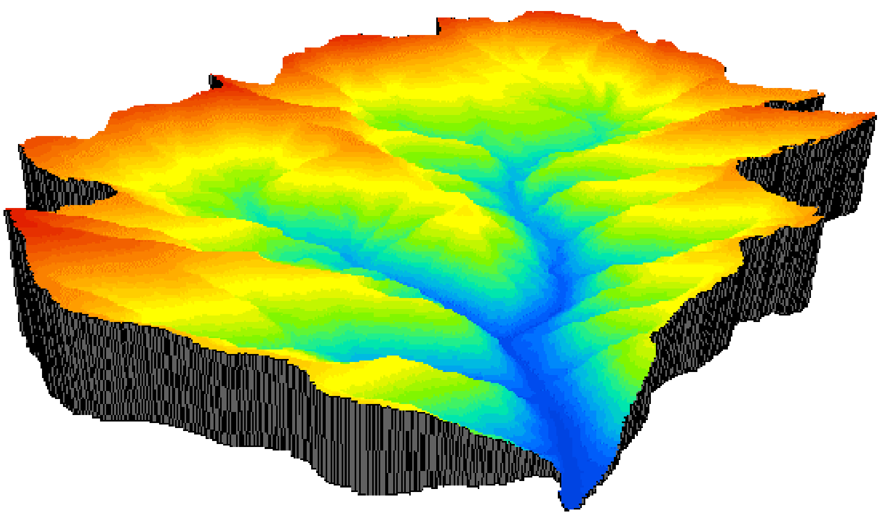


图 11-13：曲面三维透视图

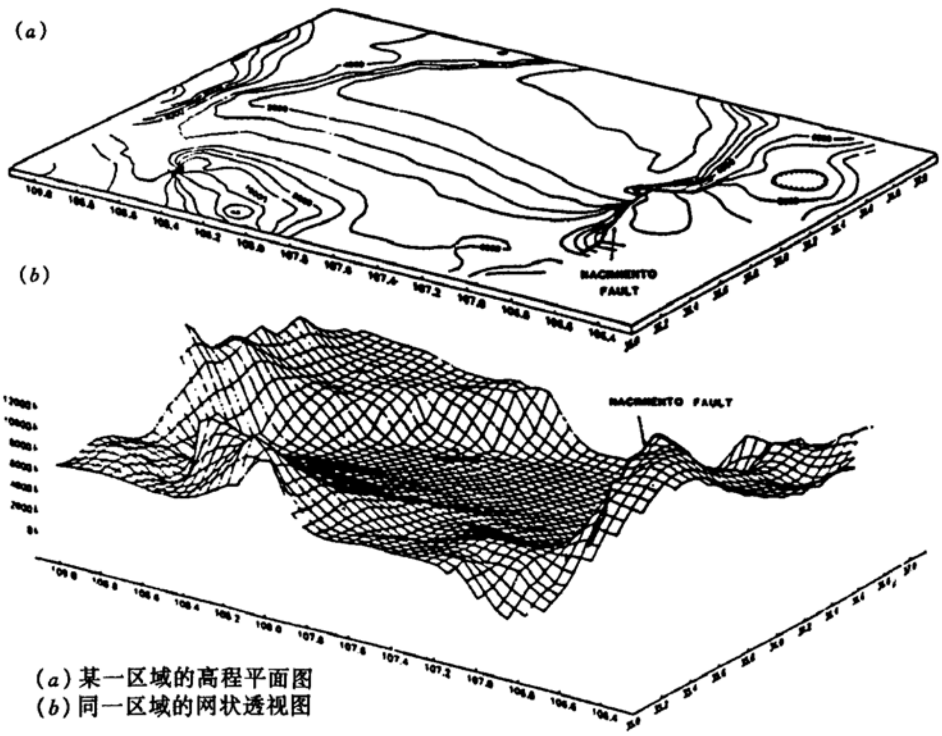


图 11-14：网状三维透视图

在利用网格 DEM 生成三维曲面以及地貌晕渲图的过程中，值得注意的一点是：如果分辨率高，则数据量很大，降低了计算速度；分辨率过低，则会使计算结果变得粗糙，这时可以考虑在网格点之间利用曲面函数拟合，但是由于曲面函数是连续的，使得结果显得“不真实”，一个较好解决办法是利用分形内插，后者可以生成细致的、更为自然的地形。

3．3 专题地图内容的表现手段

专题地图多种多样的表示方法，需要通过一定的手段来实现。选择合理的表示方法和表现手段，是提高科学内容表现能力的保证。现有的表达手段可归纳为色彩、线条和注记三种。

3．3．1 色彩

世界上的物体，均有形状和色彩两个基本特征。色彩通常分为两类：一类是黑白及各种灰色，称为非彩色；一类是彩色，包括除黑、白、灰以外的颜色。

色彩不仅能弥补单色图的缺陷，丰富图幅内容，提高专题地图的使用价值，更能如实反映制图物体的自然面貌，扩大地图内容的感染力，增强地图内容的清晰易读性。图 11-15 为 RGB 彩色立方体，红绿蓝三原色被设置在彩色立方体的三个角上。立方体下方的角，三原色强度为 0，颜色为黑；横穿立方体的对角线是强度轴，范围从起点黑（R，G，B）=（0，0，0）到白（R，G，B）=（100，100，100），这里，强度以百分比来表示；立方体中任何一点的颜色都可以通过红、绿、蓝（R，G，B）三轴上的设值来描述；如（50，50，50）是暗灰色，（0，0，100）是纯蓝色等等。红、绿、蓝被称为加色，因为新的颜色都可以通过

它们加到黑色上来获得。同样，减色三原色也可以在彩色立方体予以表示，它们是青色、品红色和黄色（C，M，Y），分布在剩余的三个角上。青由蓝和绿合成，品红色由蓝和红合成，黄由绿和红合成，染料和打印油墨用减色三原色，通过从白色中减去它们的补色来生产其他颜色。这样，在影像显示时，颜色用 R、G、B 来度量，在打印和绘图时，则用 C、M、Y 来度量。

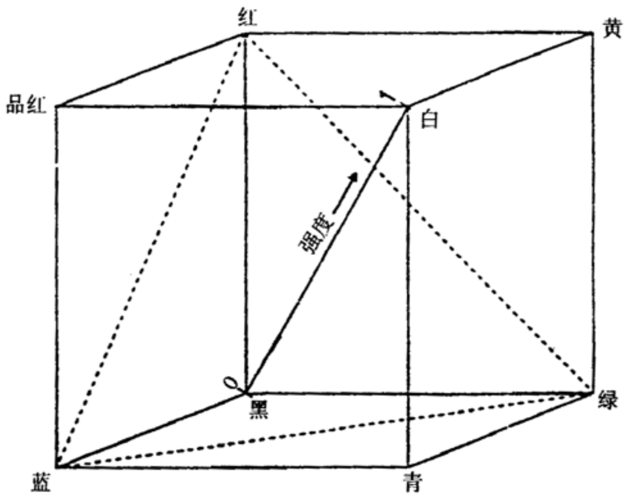


图 11-15：RGB 颜色立方体

在实际印刷出版应用中，由于难以用青色、品红色和黄色合成真正的黑色，因此一般采用的是 C、M、Y、K（黑色）色彩度量。除了 RGB 和 CMYK 颜色度量空间之外，还有其它一些颜色表述方式：

1）HLS 颜色表述，HLS 指色调（Hue），亮度（Brightness）和饱和度（Saturation），它更能够反映人对色彩的感知。RGB 色彩空间尽管组成了一个立方体，但是并非是线性的，即两个颜色的相近程度不能简单的用欧式距离来度量。而 HLS 色度空间则较好的解决了该问题，色调反映了人们对颜色的分类，如“纯红”、“品红”、“桃红”等属于红色调；亮度说明了颜色“黑”的程度，白色亮度最高，黑色最低；饱和度则反映了颜色的纯度，如果一个颜色掺入灰色，则饱和度降低，例如，桃红色饱和度低于纯红色。

2）CIE 颜色系统，CIE（Commission International de I’Eclairage）颜色系统是国际的通用颜色标准。它提供了一个一致的颜色定义方法，但是过于复杂，不易掌握。

3）孟塞尔[Munsell]颜色系统，这是被广泛使用的进行色彩设计时的系统，它采用色调、色值（Value 相当于亮度）、色品（Chroma，相当于饱和度）来描述颜色，并且使用一个比例数值来反映人的感知。

各种色彩度量分别应用于不同的方面，在地理信息系统中，因为同时要考虑屏幕显示和制图输出，所以通常要兼顾考虑 RGB 色度和 CMYK 色度。各种色度系统之间可以相互转换，转换一般采用经验公式。一个好的转换公式可以使彩色图像在不同的输出设备上达到一致的视觉效果，但是往往较为复杂。

3．3．2 线条

线条实质上是点、线、面状符号的简称，是显示制图现象数量和质量的重要手段，与色彩搭配使用，可大大地增加科学内容的表现能力。如图 11-16 所示。

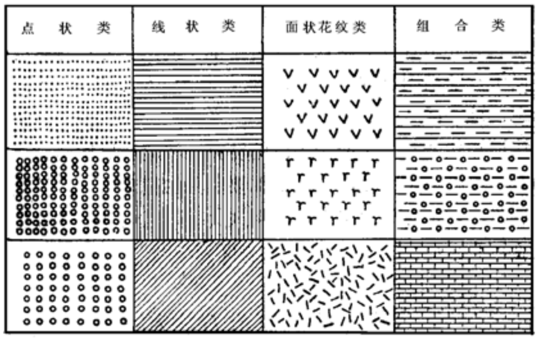


图 11-16：各种线条样式

3．3．3 注记

注记是上述两种手段的补充，注记应紧密结合专题内容，可增强现象的显现效果。对于类型图和区划图，一般都要利用注记。

GIS 中使用的硬拷贝设备：

在 GIS 制图输出时，要采用硬拷贝设备（Hard Copy Devices），所谓硬拷贝设备，是相对于屏幕显示的“软拷贝”而言，前者依赖于数据，并且不能长期保存，而硬拷贝输出的内容则可以长期保存。

1．栅格打印机和绘图仪

栅格输出设备通过点阵产生的图像来输出地图，根据输出质量高低、输出幅面大小的不同，又可以分为：

1．1）点阵打印机

点阵打印机是最简单的栅格打印机，适用于对输出质量要求不高、节约成本的情形。

1．2）喷墨打印机

与普通点阵打印机相比，喷墨绘图仪可以打印幅面更大、质量更高的地图，目前高质量的喷墨绘图仪输出精度可以超过 1000DPI。

1．3）静电绘图仪

主要用于输出大幅面（超过 2m）的地图。

1．4）激光打印机

目前黑白激光打印机主要用于输出文本，也可用来打印灰度图形，彩色激光打印机可以用于打印高质量的地图。

1．5）激光绘图仪

激光绘图仪可以达到非常高的分辨率（2000DPI），用于打印大幅面地图。

1．6）热蜡（热转印）打印机

热蜡打印机可以打印更亮、对比度更强的图像。

1．7）热升华打印机

热升华打印机机理是将颜料加热升华为气态，并输出到纸面上，热升华使得每个输出稍微模糊，适合于输出照片、图像等，但是不适于打印文字。

2．矢量模式的笔式绘图仪

笔式绘图仪直接以矢量方式工作，它通过一系列的命令来操纵绘图笔进行绘制，它可以达到很高的输出分辨率，但是实现时需要将地图数据全部转换为矢量格式，笔式绘图仪的另一个不足是输出复杂的地图时，需要时间较长。

3．照相

即用照相机对屏幕进行拍摄，得到图像，这种方式最为简单，在精度要求不高的情况下可以采用。

4．专题地图设计

专题地图的总体设计，是指任务和要求明确后初步提出的图幅基本轮廓，包括投影选择、明确比例尺、划定图幅范围、进行图面规划和绘制设计略图等内容。

4．1 图幅基本轮廓的设计

专题地图的总体设计，比普通地图和国家基本地形图复杂多样。编制一幅专题地图，不仅要学科专业与制图的紧密结合，而且要对图幅的用途和使用者的要求有深入的了解和掌握。在此基础上，才能进行设计图幅的基本轮廓。具体要了解的内容包括：

1）该图幅是专用还是多用

专题地图既能专用也可多用，而且愈来愈向多用方向发展，并相应地产生了一版地图多种式样的作法。

2）已出版的类似专题地图

分析这些图件在使用中的优缺点，吸收长处，改进不足，以便更好地满足地图使用者的需要。

3）明确地图使用者的特殊要求

根据不同的读者对象、不同用途以及不同使用场合等要求，考虑到所编制的专题地图是作为规划用的还是作参考用的、教学用的等，予以满足。

在弄清上述图幅的用途与要求之后，就要明确总体设计的指导思想，拟定专题内容项目，突出重点，提出图幅总体设计的方案。

4．2 制图区域范围的确定

专题地图图幅的区域范围，是根据用途和要求来确定的。范围选择是否合适，在很大程度上影响着图幅的使用效能，并与专题地图的数学基础有紧密的联系。与普通地图一样，根据图幅范围可分为单幅、单幅图的“内分幅”、分幅三种形式。

4．2．1 单幅

这是指一幅图的范围能完整地包括专题区域而言，通常叫截幅。专题区域放置在图幅的正中，它的形状确定了图幅的横放、竖放和长宽尺寸。专题区域与周围地区的关系要正确地处理。为了便于阅读和使用，专题地图一般以横放为主要式样；有些专题区域性形状是长的，而地图的方向习惯是上北下南，所以只好竖放。

4．2．2 单幅图的“内分幅”

这是指超过一张全开纸尺寸而分为若干印张而言。“内分幅”应按纸张规格，一般分幅不宜过于零碎，分幅面积的大体相同。

4．2．3 分幅

分幅是地形图普遍采用的一种形式。分幅图不受比例尺限制，分幅图的分幅线是根据区域大小采用矩形分幅和经纬线分幅的，分幅图原则上是不重叠的。

此外，图廓内专题区域以外的范围如何确定，在总体设计时也应明确下来。方法有：

第一，突出专题区域线，区内区外表示方法相同，只把专题区域界线加粗，或加彩色晕边，以显示专题区域范围，同时也能与相邻区域紧密联系；

第二，只表示专题区域范围，区域外空白，突出专题区域内容，区内要素与区外没有什么联系。

第三，内外有别，即专题区域内用彩色，区外用单色，且内容从简。这是专题地图普遍采用的方法。

4．3 专题地图数学基础的设计

专题地图数学基础包括地图投影、比例尺、坐标网、地图配置与定向、分幅编号和大地控制基础等，其中地图投影和比例尺是最主要的。

4．3．1 影响数学基础设计的因素

1）专题地图的用途与要求

这是影响数学基础设计的主导因素，因为投影和比例尺都是根据图幅的用途和要求选择设计的。

2）制图区域的地理位置、形状和大小

该要素是一个重要的因素，位置和形状往往影响投影和比例尺的选择。在设计时对制图区域的形状和大小要详细研究，并同时设计几个方案，选择一个合理的方案。

3）地图的幅面及形式

地图的幅面及形式都对数学基础设计有一定的影响，直接关系到使用效果。

4．3．2 投影和比例尺的设计

1）投影设计

在专题地图制图中采用较多的是等积投影和等角投影，具体设计时采用何种投影，要视专题地图的用途和要求而定。

2）比例尺设计

专题地图比例尺的设计应考虑图幅的用途和要求，根据制图区域形状、大小，充分利用纸张有效面积，并将比例尺数值凑为整数。在实际设计地图比例尺的工作中，往往还会出现一些特殊的问题，如：不要图框或破图框、移图、斜放。

4．4 图面设计

专题地图不仅要有科学性，而且也要有艺术性。图面设计包括图名、比例尺、图例、插图（或附图）、文字说明和图廓整饰等。

4．4．1 图名

专题地图的图名要求简明图幅的主题，一般安放在图幅上方中央。字体要与图幅大小相称，以等线体或美术体为主。

4．4．2 比例尺

比例尺有两种表示方法：一是用文字（如一比四百万）或数字（如 1：4000000）表示；二是用图解比例尺表示。图解比例尺间隔也有两种划分方法，一种是按单位长度划分，表明代表的实际长度；一种是按实地公里数划分，每格是按比例计算在图上的长度。比例尺一般放在图例的下方，也可放置在图廓外下方中央或图廓内上方图名下处。

4．4．3 图例

图例符号是专题内容的表现形式，图例中符号的内容、尺寸和色彩应与图内一致，多半放在图的下方。

4．4．4 附图

附图是指主图外加绘的图件，在专题地图中，它的作用主要是补充主图的不足。专题地图中的附图，包括重点地区扩大图、内容补充图、主图位置示意图、图表等。附图放置的位置应灵活。

4．4．5 文字说明

专题地图的文字说明和统计数字，要求简单扼要，一般安排在图例中或图中空隙处。其他有关的附注也应包括在文字说明中。

专题地图的总体设计，一定要视制图区域形状、图面尺寸、图例和文字说明、附图及图名等多方面内容和因素具体灵活运用，使整个图面生动，可获得更多的信息。图 11-17 展示了一些不同风格的图面设计。

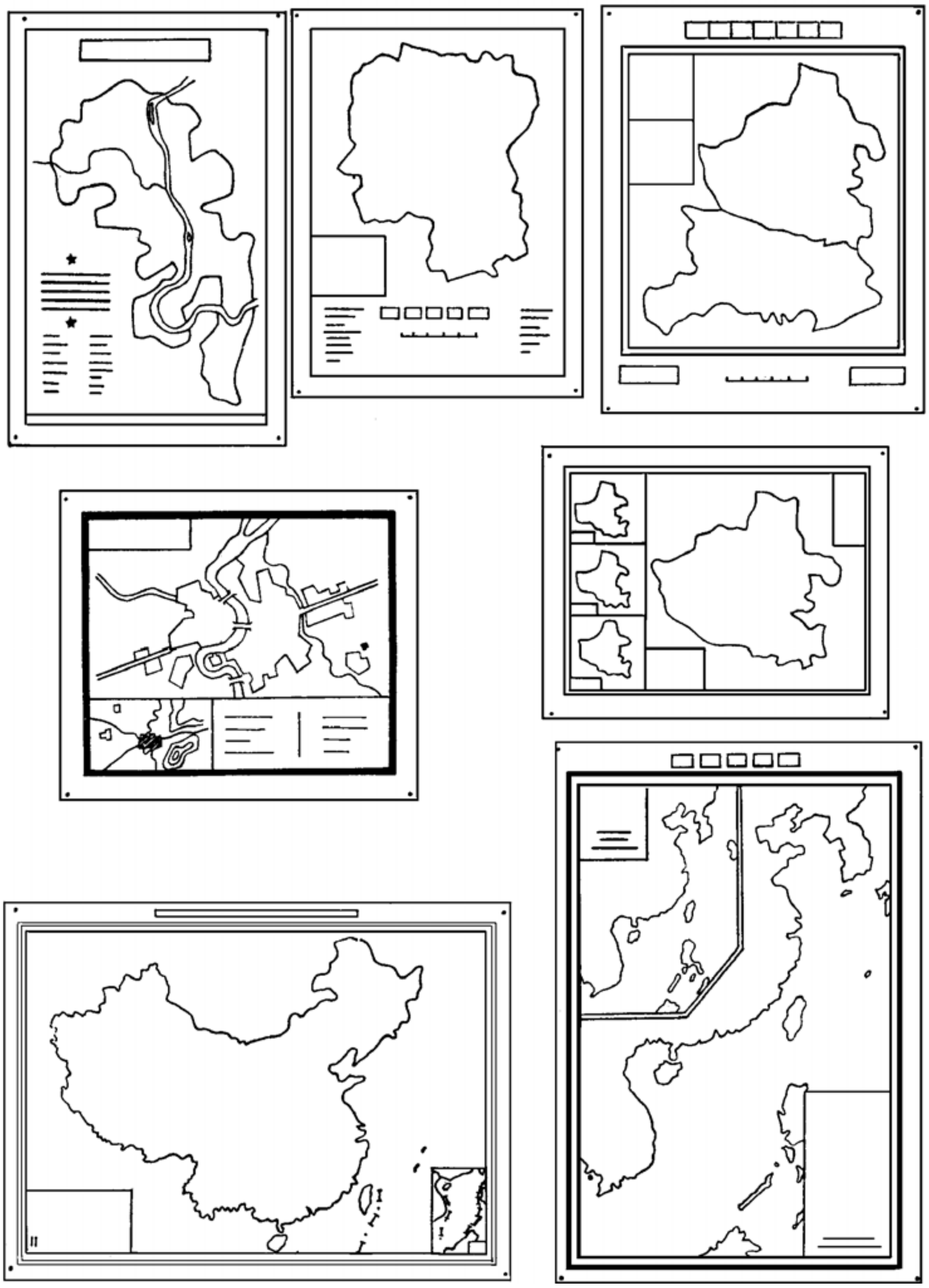


图 11-17：图面设计示例

5．制图综合(制图概括)

地理信息系统中地理信息综合功能的研究和开发是当前 GIS 进步与应用发展所面临的一个新的挑战。由于 GIS 技术和应用迅速发展所处理的问题也更加复杂化，这使得 GIS 中地理信息的有效利用和表达问题显得越来越突出。因此，人们不得不开始关注 GIS 中地理信息综合功能，将基本比例尺数据派生出多比例尺、多用途和多专题的更多有用信息，并提高地理信息的表达和利用的效率，满足 GIS 的可视化和各种分析、制图输出的需要。

5．1 制图综合的概念

制图综合是对制图区域客观事物的取舍和简化。经过概括后的地图可以显示出主要的事物和本质的特征。地图的比例尺、用途和主题，制图区域的地理特征以及符号的图形尺寸是影响制图综合的主要因素，制图综合主要表现在内容的取舍、数量简化、质量化简和形状化简等方面。

专题地图制图综合的实质，就是在科学分析研究各种专题内容要素特征的基础上，根据图幅的用途和比例尺，将图幅的专题内容加以概括化，把最主要的要素、对象的基本轮廓、主要的特征和基本规律反映在地图上。专题地图制图综合围绕着几何性和地理性两方面进行。在几何性方面对制图对象形状和大小综合简化，在地理性方面对专题内容的综合概括。制图综合就是几何正确性和地理真实性的紧密结合。

5．2 影响制图综合的主要因素

制图综合的程度受各种因素的影响，主要因素有：地图比例尺、地图的主题和用途、制图区域的地理特征以及符号的图形尺寸等。

5．2．1 地图比例尺

比例尺对制图综合的影响非常明显。由于比例尺的缩小，同一个制图区域在图上的面积随之缩小，因而图上所能表示的地物数量也相应减少。当地图幅面一定时，不同比例尺地图所包括的实地范围不同，大比例尺地图所包括的地面面积较小，小比例尺地图包括的地面面积较大。在不同范围内，对同一地物重要程度的评价并不相同，有些事物从小范围看是重要的，但在大范围内可能是次要的。

5．2．2 地图的用途

任何一幅地图所能表示的内容都是有限的，只能满足某一方面或某几方面的要求，所以地图内容的选择和表示，就必须考虑到地图的用途。地图用途既然直接决定着地图内容，因而就影响着制图综合。例如 1：50 万比例尺地形图和 1：50 万比例尺政区图，由于两者用途不同，地图内容的取舍和化简程度就很不一样。1：50 万比例尺地形图是国家基本地形图之一，在军事上是战役用图，经济上是规划用图，在科学研究方面是参考用图，在这种地图上要全面地反映制图区域的自然和社会经济方面的基本情况；而 1：50 万比例尺政区图是省级机关进行行政管理、布署工作用图，在图上主要表示社会经济要素（如境界、居民地、交通线等），而对自然要素（如地形、土质、植被等）则可作较大的概括。

5．2．3 地图的主题

地图的主题即地图上所反映的主要内容，它决定各事物在图上的重要性，因而影响制图综合程度。这在专题地图上表现特别明显，例如比例尺相同的同样地区但主题不同的两幅地图，一幅是地形图，一幅是政区交通图，前者主要表示地形特征及水系，居民地较少，不表示铁路和公路；后者重点表示行政区划、铁路和主要公路、各级行政中心以及作为交通枢纽

的居民地，对于水系则只是表示主要的河流和湖泊，一般不表示地形。

5．2．4 制图区域的地理特征

制图区域地理特征是指该区域的自然和社会经济条件，地面上不同地区具有不同的地理特征，因此制图综合时就要选取那些能反映区域特征的事物，舍去那些不代表区域特征的事物。同样的地理事物在不同地区具有不同的意义，这就影响着制图综合。例如，小城镇在人口稠密的地区是次要的，图上一般不表示，而在人口稀少的地区，小城镇则成为主要的居民地，图上要选取。

5．2．5 符号的图形尺寸

各种地理事物，在图上均以符号表示。符号的图形、尺寸直接关系着地图的负载量，因而也影响着制图综合的程度。图形尺寸规定得小一些，选取的内容就可以多一些，概括程度就小一些，地图的内容就会详细一些，图面负载量就大一些；反之，地图内容必然简略，图面负载量就小。此外，几何图形的最小尺寸与图形的结构和复杂程度有关，轮廓符号的最小尺寸受轮廓界限的形式、内部颜色和背景等一些因素影响，可见符号图形最小尺寸是确定制图综合的必要参考数据。

5．2．6 可视化要求

可视化是地理信息处理的窗口与处理结果的直观表达形式，因而是决策的直观依据。只有把空间数据库中的海量数据转换为直观的图形信息，地理信息处理结果才能为规划、管理与决策提供有力的支撑。

5．3 制图综合的基本方法

制图综合是一个高度智能化的和具有创造性的作业过程，它是一个整体任务，包含了一系列不同性质的操作，可以分解为若干个子过程来实现。由于这些子过程之间缺少明确的内在或逻辑联系，使得这些子过程是以某种混合形式来组合应用。不同学者对制图综合过程提出了不同的分解模式：如三算子模式：选取、概括和移位；四算子模式：选取、概括、合并和移位；七算子模式：纯几何综合包括简化、夸大、位移；几何/概念综合包括选取、合并类型化和强调。

传统的制图综合包括的内容主要有：选择主要的对象，确定主要对象的选取标准；概括数量特征，主要从缩减分级、改变表示方法着手；综合质量特征，将一些相近的分类分级予以归并，利用组合符号代替单个符号；简化轮廓图形，包括简化符号线划和轮廓界线，合并小面积图斑成为大面积的轮廓，夸大具有特征的小面积轮廓；去掉一些小指标，转化为总的指标等。

启发式制图综合是把整个优化过程分解为若干个子过程来实现，这些子过程基本上等同于传统制图综合的某些手法（如选取、概括、合并与位移等）。在概括过程中，如弯曲海岸线的概括意味着相应的港湾或海角的取舍。“英国海岸线长度问题”是新兴学科“分形学”中的一个典型的问题。分形几何是用来描述难以用欧式几何来描述的具有多层嵌套的自相似

结构。地图中的河系，地貌中的沟叉，城市中的大道、胡同、小巷、里弄等，均具有嵌套特征，是典型的分形结构。分形几何原理在地图信息处理中必将得到应有的深入应用。

分形：

分形一词（Fractal）是 B. Mandelbrot 用拉丁词根创造的单词，意思是细片、破碎、分数、分级等。它是那些局部和整体按照某种方式相似的集合[B. Mandelbrot]。分形具有以下特征：

1）具有精细的结构，在任意小的尺度之下，总有复杂的细节；

2）分形是不规整的，它的整体和局部都不能用传统的几何语言来描述；

3）通常具有自相似形式，这种自相似可以是近似的或者是统计意义上的；

4）一般的，分形的某种定义之下的分形维数大于其拓扑维数，例如在平面上，一条分形曲线的维数在 1-2 之间；

5）在大多数情况下，分形以非常简单的方法确定，可以由迭代过程产生。

分形是描述不规则几何形态的有力工具，而自然界的许多物体，都具有分形的特性，如海岸线、地形、云、甚至生物等等。对于一条海岸线而言，无论从遥感卫星上看，还是从岸边的悬崖上看，它总是曲折的。因此，对于这些地物，要完全“精确”的描述其几何形状是不可能的，在 GIS 数字化的过程中，要进行抽样操作。

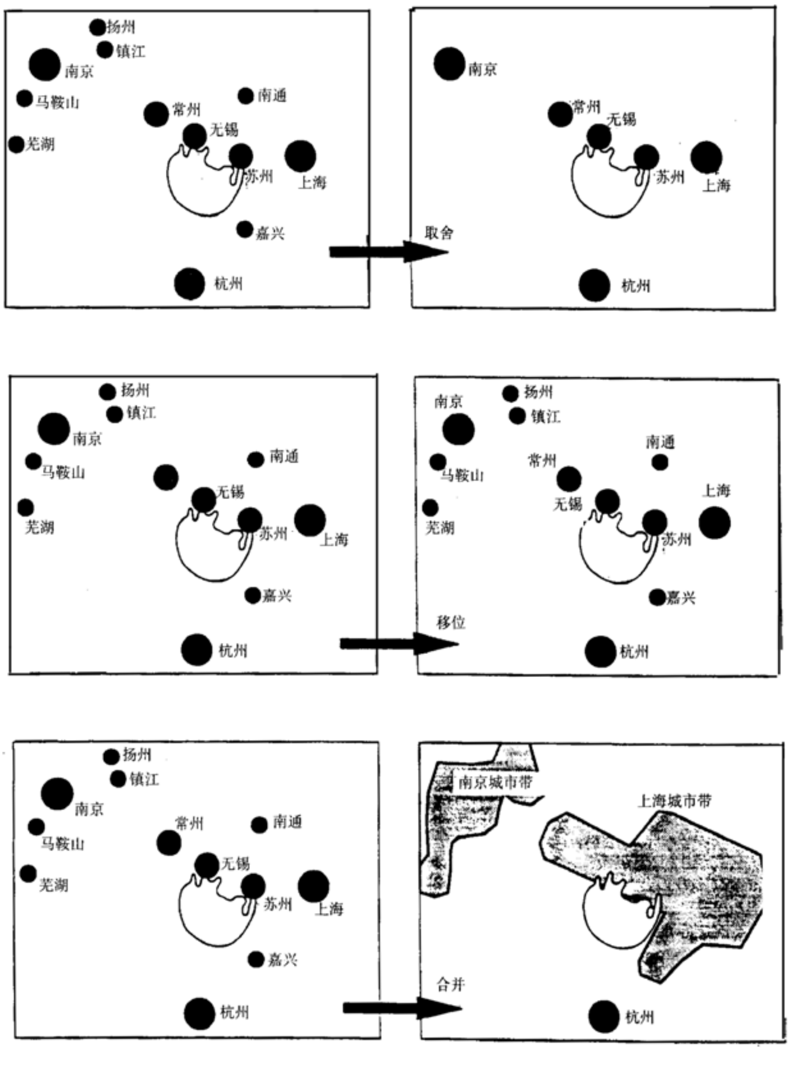


图 11-18：点符号的综合

总之，制图综合的基本方法为内容的取舍、数量化简、质量化简和形状化简。图 11-18、

图 11-19 和图 11-20 分别是对点状符号、线状符号和面状符号的综合。

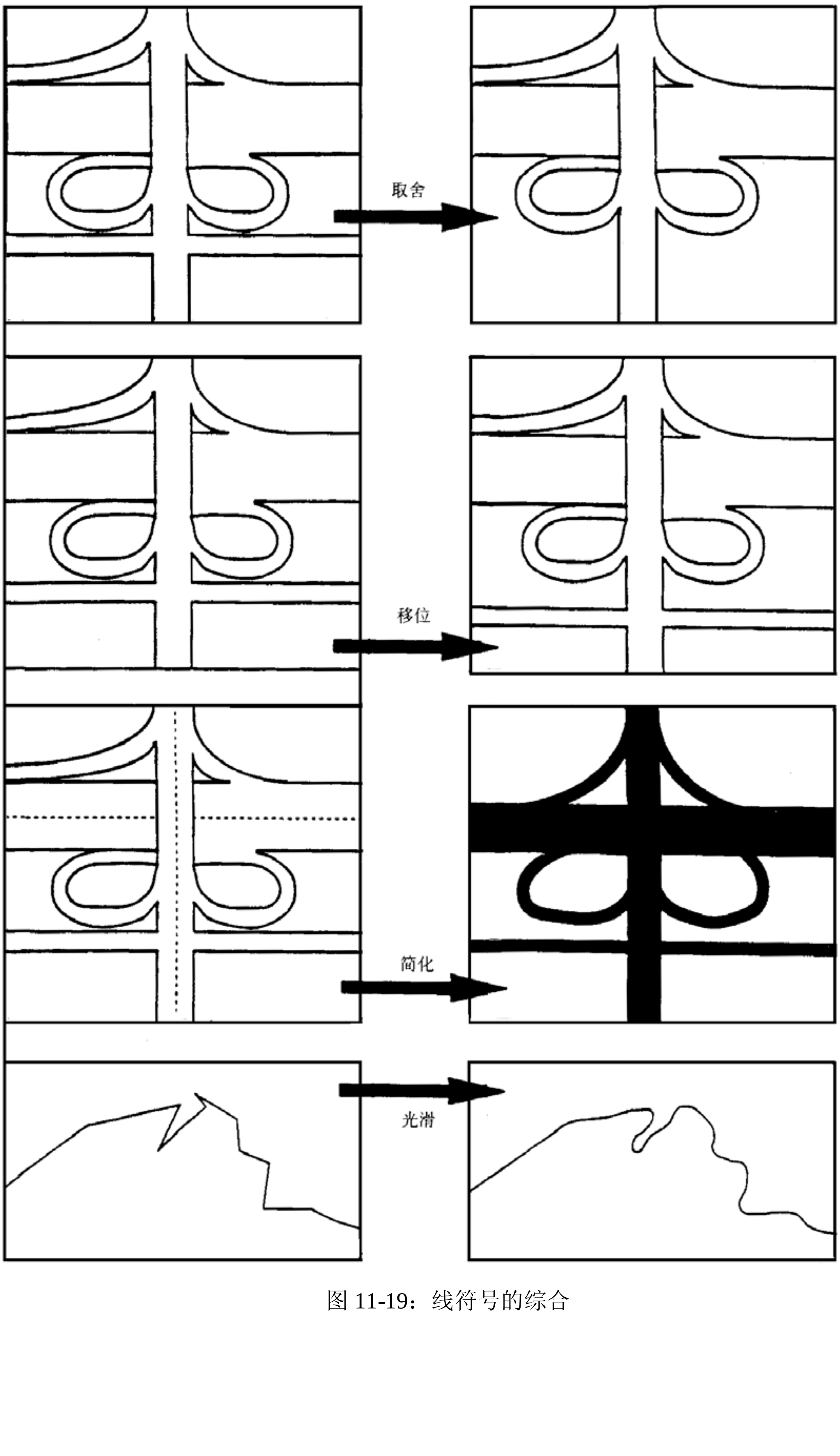


图 11-19：线符号的综合

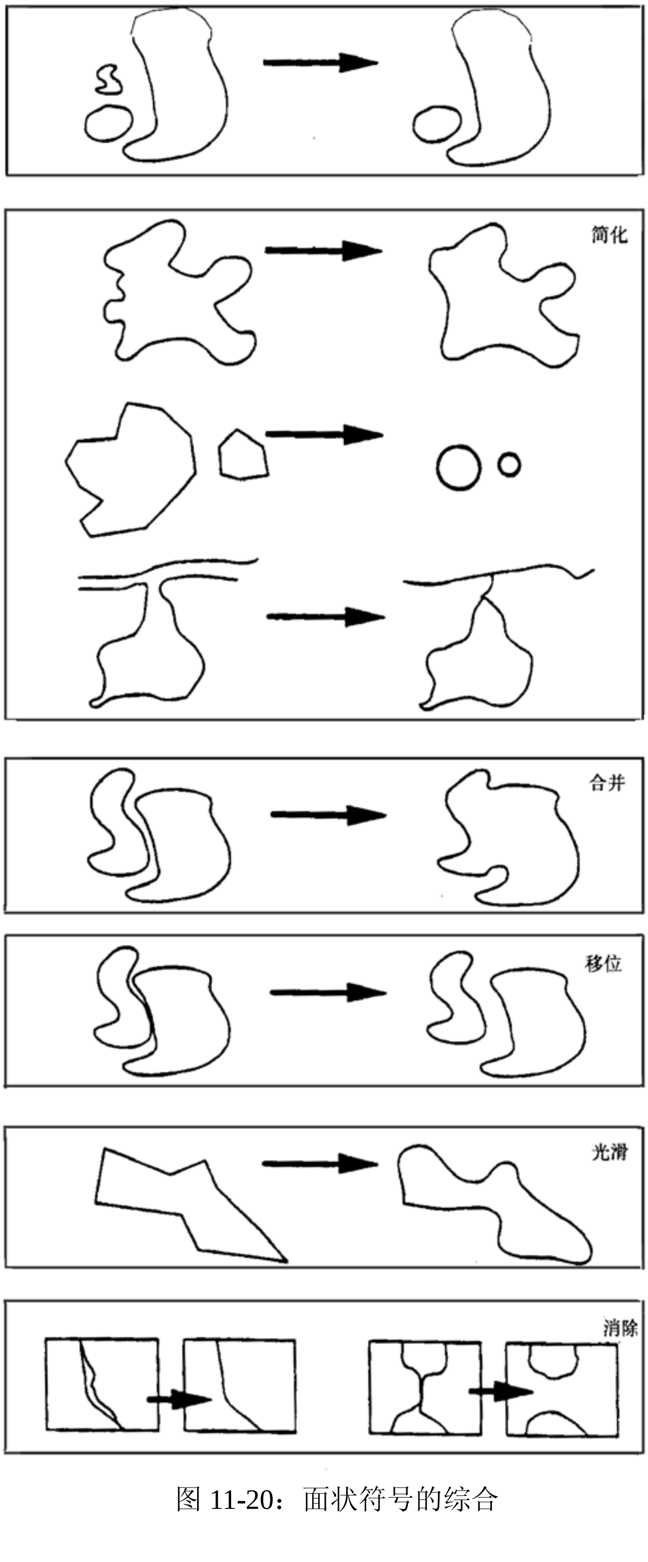


图 11-20：面状符号的综合

5．3．1 内容的取舍

内容的取舍是指选取较大的、主要的内容而舍去较小的、次要的或与地图主题无关的内容。选取主要表现在：一是选取主要的类别，二是选取主要类别中的主要事物。舍去主要表现在：一是舍去次要的类别，二是舍去已选取的类别中的次要事物。注意所谓的主要与次要

是相对的，它随着地图的主题、用途、比例尺的不同而异。地图内容的选取，一般按下列顺序进行：

1）从整体到局部：进行选取时，要首先从整体着眼，然后从局部入手。

2）从主要到次要：地图上所表示的各个要素，根据地图的主题和用途，总有主要与次要之分。选取时要遵循先主要后次要的顺序进行。

3）从高级到低级：这样可以保证较高级的能被选择，不至于遗漏。

4）从大到小：这样可以保证大的事物首先入选。

总之，选取时要从总体出发，首先选取主要的、高级的、大型的事物，再依次选取次要的、低级的、小型的事物，最后还要从整体上进行分析，观察是否反映了制图区域的总体特征。

5．3．2 质量特征的化简

地图上各事物的质量差别通常是以分类来体现的。质量特征的化简就是减少一定范围内事物的质量差别，用概括的分类代替详细的分类，即按事物的性质合并类型或等级相近的事物。例如，将针叶林、阔叶林和混交林合并为森林；将喀斯特山地、喀斯特丘陵、喀斯特台地、喀斯特溶蚀堆积盆地合并为喀斯特地貌。

5．3．3 数量特征的化简

数量特征的化简就是减少事物的数量差别，增大数量指标内部变化的间距，对于数量指标低于规定等级的事物不予表示。在用等值线表示数量特征的地图上，化简时要扩大等值线间距值。进行数量特征化简时，不仅要考虑地图比例尺和用途，而且要特别注意考虑事物数量分布的特点及保持具有质量意义的分级界限。

5．3．4 形状化简

形状化简用于呈线状和面状分布的事物。形状化简的目的在于保留事物本身所固有的、典型的特征。化简的方法有删除、夸大和合并。删除就是去掉那些因比例尺缩小而无法清楚表示的碎部。有时为了显示和强调事物平面图形的特征，将本来按比例应删除的小弯曲，夸大表示出来。合并就是将邻近的、间隔小到难以区分的同类事物的图形加以合并，以表示出事物的总体特征。对某种事物进行形状简化时，要考虑到与其他事物的关系，使彼此之间能协调一致。例如对湖泊进行形状化简时，要注意与地形、水系的关系。

6．地理信息的可视化

6．1 基本概念

可视化(Visualization)是指在人脑中形成对某物（某人）的图像，是一个心理处理过程，促使对事物的观察力及建立概念等。科学计算可视化是通过研制计算机工具、技术和系统，把实验或数值计算获得的大量抽象数据转换为人的视觉可以直接感受的计算机图形图

像，从而可进行数据探索和分析。把地学数据转换成可视的图形这一工作对地学专家而言并不新鲜。测绘学家的地形图测绘编制，地理学家、地质学家使用的图解，地图学家专题、综合制图等，都是用图形(地图)来表达对地理世界现象与规律的认识和理解。科学计算可视化与上述经典常规工作的最大区别是科学计算可视化是基于计算机开发的工具、技术和系统，而过去地学中的可视表达和分析是手工或机助的(计算机辅助制图)，并把纸质材料作为地图信息存储传输的媒介。

科学计算可视化，自从 80 年代末提出以后，得到了迅速的发展并成为一个新兴的学科，其理论和技术对地学信息可视表达、分析的研究与实践产生了很大的影响。国际地图学会(ICA)在 1995 年成立了一个新的可视化委员会，并在 1996 年 6 月与计算机器图形协会(ACM SIGGRAPH)合作，开始一个名为“Carto-Project”的研究项目，其目的是探索计算机图形学的技术与方法如何更有效地应用在地图学与空间数据分析方面，促进科学计算可视化与地图可视化的连接和交流。地学专家对可视化在地学中的地位和作用，已进行了比较深入的讨论，从不同的角度提出了与可视化密切相关的地图可视化、地理可视化、GIS 可视化、探析地图学(Exploratory Cartography)、地学多维图解、虚拟地理环境等概念，但有不同的理解，对其相互关系的认识也不明确。

地理信息系统的多维可视化是指采用 2.5 维、三维和四维等地图表现形式来反映地理客体的多维特征，其中 2.5 维形式是图面上有隐藏部位的鸟瞰式地图表现形式，又称“假三维”，例如表示矿床的面层，可用显示为同分异状的等值线或不规则三角网中的小块平面来表示，而面上的高程值都不是一个独立的变量，在任一给定的位置仅能用一个高程值表示一个面。三维形式则是真正的三维立体显示，三维技术可分为基于面表示和基于体表示，其中面表示可容易地为地层及其构造提供精确的空间描述，特别是构造复杂的地带或岩石断层处，然而这些面仅将地表面以下部分分成许多区域，各区域内的不同地质条件不容易描述；体表示将整体细分成大量的体元（Voxels），体元及其变异的一个主要的优点是易于表示异质特征的整个三维分布状况。四维是除了三维立体以外，再增加一维属性值（一般是时间维），GIS 的时间查询能力目前还不常见，但是地理学家可能想对某一时刻的所有地理条件或某一时间段内的平均地理条件进行评价，这时，就需要获得不同时刻的数值。为了充分满足这种需要，时间数据获取能力应该与三维模型相结合，如分析并预报水灾、地震、暴风雨及滑坡等。

地理信息的多维可视化在地球科学中具有重要意义，它对于动态地、形象地、多视角地、全方位地、多层面地描述客观现实，对于虚拟化研究、再现和预测地学现象，都有突出的方法论意义。例如，在地质科学中用真三维反映地下矿体、矿脉（如含油体、含水体、金属矿脉等），能够帮助人们发现用常规手段难以发现的地质现象和矿藏；在大气科学中用四维（真三维加时间维）形式表示气旋、龙卷风、降水云系的发生、发展和演化过程；在地理学中用四维方式模拟整个河床内洪水的流动、涨落、对河堤的侵蚀，以及决堤后封决口时的水下状况，等等，都具有十分重要的科学价值和明显的实用意义。

6．2 地学可视化的类型

地学可视化包括地图可视化、地理信息系统(GIS)可视化及其在专业应用领域的可视化。地图可视化与 GIS 可视化是地学可视化理论和技术的两个基础部分。地学多维图解是从非可视化角度提出的，关于图的地学知识表达、获取以及图的地学问题解的求证原理和方法，是地学可视化的高一级发展阶段。地图可视化中的信息表达交流模型和地理视觉认知决策模型的研究，将从理论上促进地学多维图解模型的进一步发展。基于虚拟现实技术和计算机网络技术的虚拟地理环境是地学可视化、地学多维图解理论和技术发展的最后集成系统。

地学可视化是科学计算可视化与地球科学结合而形成的概念，是关于地学数据的视觉表

达与分析，下面探讨该领域中的三个概念：地图可视化、地理可视化和 GIS 的可视化，及其相互关系。

6．2．1 地图可视化

可视化对现代地图学的发展有着极其重要的作用，虽然对其理解还有着一定的分歧。Taylor 强调了计算机技术基础支持下的地图可视化，并认为可视化包括交流与认知分析。MacEachren 则强调交流与可视化在地图学中的应用，而不是地图的技术制作，并把可视化与交流作为并列的两个要素。

在可视化的过程中，信息交流传输以及认知分析的界限并不需要很清楚地划分。事实上，它们一直融合在一起，并具有共同作用的特点。由于可视化的两个特点，即交流与认知分析，从而对现代地图学的理论和方法产生影响。图 11-21 表示了地图可视化的概念框架。

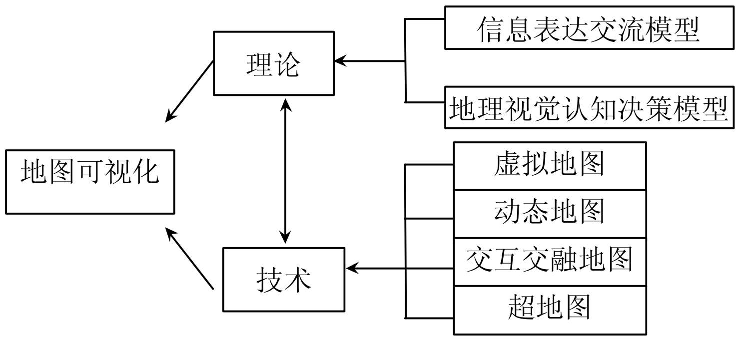


图 11-21：地图可视化的概念框架

由于可视化具有交流与认知分析的两个特点，从而使信息表达交流模型与地理视觉认知决策模型构成了地图可视化的理论，而这两个模型将应用于计算机技术支持的虚拟地图、动态地图、交互交融地图以及超地图的制作和应用等。

虚拟地图指计算机屏幕上产生的地图，或者利用双眼观看有一定重叠度的二幅相关地图，从而在人脑中构建的三维立体图像。虚拟地图具有暂时性，实物地图具有静态永久性。虚拟地图和人的心智图像相互联系与作用的原理和过程，同传统的实物地图是不一样的，需要建立新的理论和方法。

动态地图是由于地学数据存贮于计算机内存，可以动态地显示关于地学数据的不同角度的观察，不同方法(如不同颜色、符号等)的表达结果，或者地学现象随时间演变的过程等等。由于地图的动态性，地学现象的表达在时间维上展开。所以，传统的关于纸质静态地图的符号制作、符号注记等制作理论和方法在动态时不再完全适合。另外，人又是如何认知分析动态的信息流等需要进一步的探讨和深入研究。

交互交融地图是指人与地图可进行相互作用和信息交流。交互地图是人可以通过一定的途径，例如选择观察数据的角度、修改显示参数等来改变地图的显示行为。在这个过程中，屏幕地图(或双眼视觉立体地图)即虚拟地图，与应用人员人脑中相关地学知识以及直觉等形成的心智图像一直处于相互作用、相互比较、相互修改完善的信息联系和反馈状态。交融地图，是人与地图的融合程度，也就是人在虚拟地图中的投入感和沉浸感(Immersion)。地图的交互和交融性的区别可以举这样的一个例子。如你到图书馆借书，交互性表示你借书要通过图书馆的服务员，首先填写条子，然后服务员去找书，最后再给你。而交融性是你可以直接进入书库进行搜寻，并发现你要借的书。虚拟现实技术是交互交融地图的发展基础。

超地图(Hyper-maps)是基于万维网(WWW)的与地学相关的多媒体，可以让用户通过主题

和空间进行多媒体数据的导航，这与超文本的概念相对应。超地图提出了万维网上如何组织空间数据并与其它超数据(如文本、图像、声音、动画等)相联系的问题。超地图对于地图的广泛传输与使用，即对公众生活、社会决策、科学研究等产生巨大的作用，具有重要的意义。

6．2．2 地理可视化

地理可视化是地学可视化中另一个被使用的概念。MacEachren 开始时采用地图可视化，但他认为地理可视化要大于地图可视化的研究范围，如遥感图像、图表、摄影影像等中的可视化在地图可视化中并不作为重点研究的对象。所以，目前他倾向于采用地理可视化，认为它包含了所有空间显示工具，而这些所有空间显示的集合就是现代地图学中所指的可视化。

Didase[1990]针对科学可视化、数据探索分析以及地理科学的应用，提出了地理可视化的作用框架，强调了地理研究过程中地图的作用，它包括数据探索，假设形成并确证，综合合成，到最后的结果表达与呈现(图 11-22)。

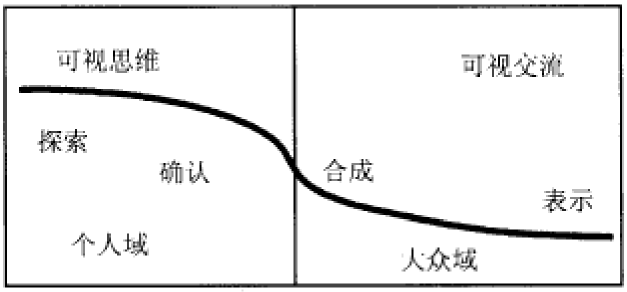


图 11-22：可视化作为地理研究工具的框架描述[Didase]

Didase 认为，可视化在研究过程的早期侧重于个人特征的视觉思维，后期侧重于研究结果的公众交流与传输，而这个特征会重新建立地图学和地理学的联系，因为在过去的二三十年中，地图学家把大部分的精力放在视觉交流传输，而地理学家(地理制图学家)在 20 世纪的前 50 年则把研究放在视觉思维与视觉分析上。

地图学与地理学作为两门经典学科，其研究对象均为区域地理系统，但前者侧重于地理空间信息的地图表达与应用，后者则把地图作为一种重要的研究工具来解决地理问题，而可视化具有的视觉交流传输和视觉分析特征，则可作为桥梁把地图学与地理学紧密地联结在一起。所以，虽然地理可视化可认为来自于可视化与地理学的结合，地图可视化来自于可视化与地图学的结合，但是由于可视化具有连结和融合地理学和地图学的特点，自然地，地图可视化与地理可视化是属于同一本质的两个概念，只是常规学科领域的划分以及研究团体的不同，导致在研究内容及范围的认识上有所侧重。所以 MacEachren 采用地理可视化，而不倾向于地图可视化的原因，仍然是从地理学与地图学领域的经典研究特征出发，而不是着眼于可视化的新技术特征，从而反映出学科领域划分对问题认识所带来的深刻影响。

6．2．3 GIS 可视化

20 世纪 60 年代发展起来的基于计算机的地理信息系统开始形成时，就利用计算机图形软硬件技术，把地理空间数据的图形显示与分析作为基本的不可缺少的功能，GIS 可视化要早于科学计算可视化的提出。GIS 可视化早期受限于计算机二维图形软硬件显示技术的发展，大量的研究放在图形显示的算法上，如画线、颜色设计、选择符号填充、图形打印等。继二维可视化研究后，进一步发展为对地学等值面(如数字高程模型)的三维图形显示技术的研究，它是通过三维到二维的坐标转换、隐藏线、面消除、阴影处理、光照模型等技术，把

三维空间数据投影显示在二维屏幕上。由于对地学数据场的表达是二维的，而不是真三维实体空间关系的描述，因此属于 2.5 维可视化。但现实世界是真三维空间的，二维 GIS 无法表达诸如地质体、矿山、海洋、大气等地学真三维数据场，所以，从 80 年代末以来，真三维 GIS 及其体可视化成为 GIS 的研究热点。随着全球变化，区域可持续发展，环境科学等的发展，时间维越来越被重视。而计算机科学的发展，如处理速度加快，处理与存贮数据的容量加大，数据库理论的发展等使得动态地处理具有复杂空间关系的大数据量成为可能，从而使得时态 GIS、时空数据模型、图形实时动态显示与反馈等的研究方兴未艾。所以，从 GIS 及其可视化的发展看，GIS 可视化着重于技术层次上，例如数据模型(空间数据模型，时空数据模型)的设计，二维、三维图形的显示，实时动态处理等，目标是用图形呈现地学处理和分析的结果。

6．3 虚拟地理环境

地学可视化的数据探索和地学多维图解的实现需要高效的人与地学数据、地学问题形成的图形场景的相互信息交流，也就是计算机系统能提供某种信息交流与反馈机制，可以让地学专家充分发挥其图形图像思维能力和想象力，可以让地学专家设想某种假设，并立即得到该假设导致的结果。虚拟现实技术可让人完全沉浸于某种计算机三维场景中，让地学专家把注意力集中于地学问题上，而其充分高效的多感觉交互能力为地学的数据探索和地学多维图解的技术实现，提供了可能。

虚拟现实(Virtual Reality)又称灵境技术，是指通过头盔式的三维立体显示器、数据手套、三维鼠标、数据衣(Data Suit)、立体声耳机等使人能完全沉浸计算机生成创造的一种特殊三维图形环境，并且人可以操作控制三维图形环境，实现特殊的目的。多感知性(视觉、听觉、触觉、运动等)，沉浸感(Immersion)，交互性(Interaction)，自主感(Autonomy)是虚拟现实技术的四个重要特征，其中自主感是指虚拟环境中物体依据物理定律动作的程度，如物体从桌面落到地面等。

虚拟现实技术、计算机网络技术与地学相结合，可产生虚拟地理环境 VGE (Virtual Geographical Environment)。虚拟地理环境是基于地学分析模型、地学工程等的虚拟现实，它是地学工作者根据观测实验、理论假设等建立起来的表达和描述地理系统的空间分布以及过程现象的虚拟信息地理世界，一个关于地理系统的虚拟实验室，它允许地学工作者按照个人的知识、假设和意愿去设计修改地学空间关系模型、地学分析模型、地学工程模型等，并直接观测交互后的结果，通过多次的循环反馈，最后获取地学规律。

虚拟地理环境特点之一是地学工作者可以进入地学数据中，有身临其境之感；另一特点是具有网络性，从而为处于不同地理位置的地学专家开展同时性的合作研究、交流与讨论提供了可能。

虚拟地理环境与地学可视化有着紧密的关系。虚拟地理环境中关于从复杂地学数据、地学模型等映射成三维图形环境的理论和技术，需要地学可视化的支持；而地学可视化的交流传输与认知分析在具有沉浸投入感的虚拟地理环境中，则更易于实现。地学可视化将集成于虚拟地理环境中。

虚拟地理环境的发展与完善，除了依赖于计算机的虚拟现实技术外，还与地学信息获取处理技术(如遥感、遥测等)，地学分析模型构建水平，地学可视化，地学专家系统，地学空间认知理论等的发展密切相关。虚拟地理环境对地学发展有重要的意义。虚拟地理学的提出就表达了虚拟地理环境对地理学未来发展的作用和影响。另外，一般认为地理科学发展缓慢的一个原因是无法进行室内试验，从而使地学假设理论无法得到实践的检验。虚拟地理环境为地学工作者提供了可重复的信息模拟实验的可能，任何一个地学分析模型均可以由其他人

在虚拟地理环境中运行模拟，受到检验，从而加速地学理论的成熟和发展。

6．4 地学可视化研究框架

可视化来源于科学计算可视化，带有浓厚的计算机技术特征。可视化与传统的地图学相互作用并融合，从而扩展了可视化本身的内涵。地学可视化除了技术特征外，还应建立起其理论范式。图 11-22 描述了地学可视化的框架，以及地图可视化、地理可视化、GIS 可视化、地学多维图解、虚拟地理环境等在框架中的位置，和它们之间的相互关系。以计算机科学为基础的地学可视化与科学可视化学科有较大程度的叠合与交叉，但并不能互相替代，两者之间的相互联系、相互交流会促进两学科的共同发展。

地学可视化包括地图可视化、地理信息系统(GIS)可视化及其它们在专业应用领域的可视化如海洋可视化、大气可视化、地质可视化、社会经济可视化等。地学可视化可从地图可视化和 GIS 可视化两方面进行理论和技术的研究。地图可视化的研究包括信息表达交流模型和地理视觉认知决策模型的构建，以及在上述模型指导下的虚拟地图、动态地图、交互交融地图、超地图的设计、制作和应用。GIS 可视化的研究包括地学体三维、时空多维数据内插加密，可视化数据模型设计， 三维、多维数据显示与分析，矢量、张量和不确定数据显示与分析，人文、经济数据可视化，实时动态交互处理，并行技术，基于网络和万维网的地学可视化，多用户合作可视化等。

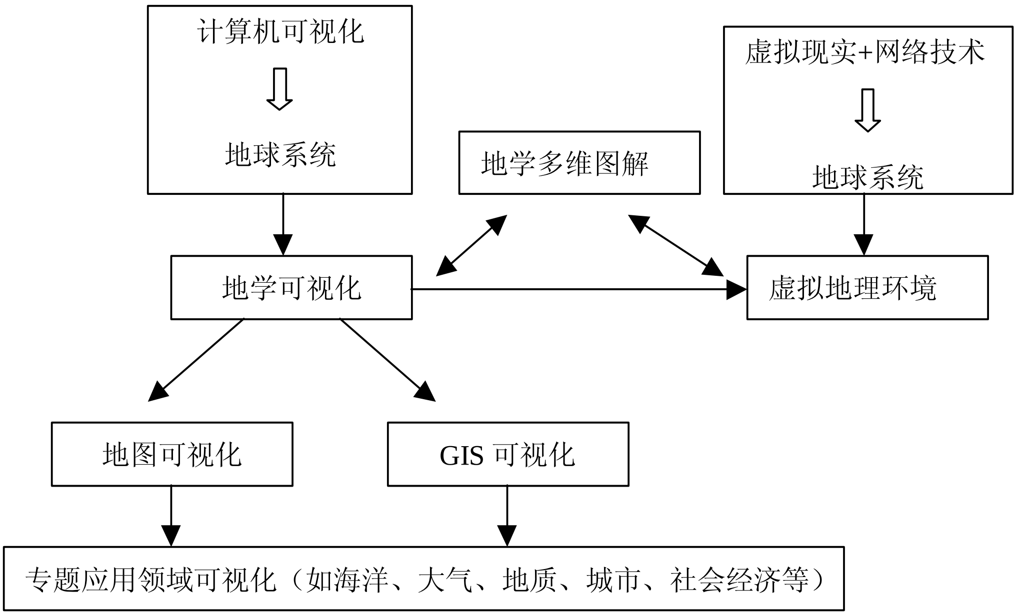


图 11-23： 关于地学可视化研究框架

地学多维图解是从非可视化角度，提出的关于图的地学知识表达、获取及其图的地学问题解求证的理论方法模式，它与地学可视化有着密切的联系，是地学可视化研究的高一级发展阶段。基于计算机的地学多维图解以地学可视化为基础，结合地学专家知识库、推理机，高效的人机交互方式等来帮助地学专家获取对某一问题的科学认识。地学多维图解的研究涉及地学图形思维，地理空间认知，图解规则获取，地学知识表达及推理，与地学可视化技术和系统的融合等。

如果把地学可视化、地学多维图解中的高效人机交互方式发展为虚拟现实技术中的融入沉浸方式以及多感觉交互方式，再加上网络连接技术等，那么就产生虚拟地理环境（VGE，Virtual Geographic Environment）。VGE 是地学可视化、地学多维图解等发展的最后集成系统，

作为一个基于计算机的地学虚拟实验室，它对地学的发展有着非常重要的意义。虚拟地理环境的研究内容包括 GIS 数据模型，三维、时空多维大数据量的实时动态交互，网络虚拟现实构模语言（VRML）及其他网络开发工具如 Java，ActiveX 等的运用等等，它们也是数字地球技术的基础。

吾尝跂而望矣，不如登高之博见也。

荀子

第十二章 3S 集成技术

导读: 本章简介了遥感和全球定位系统技术的基本概念以及它们与 GIS 的集成应用。作为实时、客观获取空间信息的新兴技术手段，遥感和全球定位系统成为地理信息系统的重要数据来源，而通过 GIS 对其获得的数据进行处理和分析，可以提取各种有用信息，以进行决策支持。

本章主要介绍了一些 3S 集成的具体应用，并简要描述了其技术实现方案。

1．遥感简介

遥感(Remote Sensing)，通常是指通过某种传感器装置，在不与研究对象直接接触的情况下，获得其特征信息，并对这些信息进行提取、加工、表达和应用的一门科学技术。

作为一个术语，遥感出现于 1962 年，而遥感技术在世界范围内迅速的发展和广泛的使用，是在 1972 年美国第一颗地球资源技术卫星(LANDSAT-1)成功发射并获取了大量的卫星图像之后。近年来，随着地理信息系统技术的发展，遥感技术与之紧密结合，发展更加迅猛。

遥感技术的基础，是通过观测电磁波，从而判读和分析地表的目标以及现象，其中利用了地物的电磁波特性，即“一切物体，由于其种类及环境条件不同，因而具有反射或辐射不同波长电磁波的特性”（图 12-1)，所以遥感也可以说是一种利用物体反射或辐射电磁波的固有特性，通过观测电磁波，识别物体以及物体存在环境条件的技术。

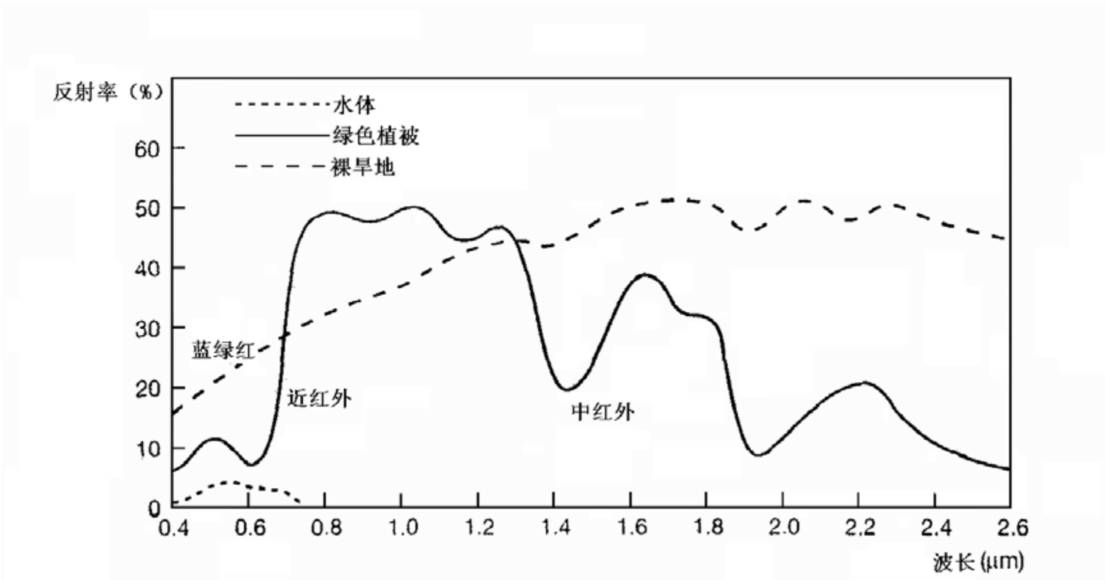


图 12-1：几种常见地物（水、绿色植被、裸旱地）的电磁波反射曲线

在遥感技术中，接收从目标反射或辐射电磁波的装置叫做遥感器（Remote Sensor），而搭载这些遥感器的移动体叫做遥感平台（Platform)，包括飞机、人造卫星等，甚至地面观测车也属于遥感平台。通常称用机载平台的为航空遥感（Aerial Remote Sensing)，而用星载平台的称为航天遥感。

按照遥感器的工作原理，可以将遥感分为被动式遥感（Passive Remote Sensing）和主动

式遥感（Active Remote Sensing）两种，而每种方式又分为扫描方式和非扫描方式，其中陆地卫星使用的 MSS（Multispectral Scanner）和 TM（Thematic Mapper）属于被动式、扫描方式的遥感器（图 12-2），而合成孔径雷达(SAR-Synthetic Aperture Radar)属于主动式、扫描方式的遥感器。

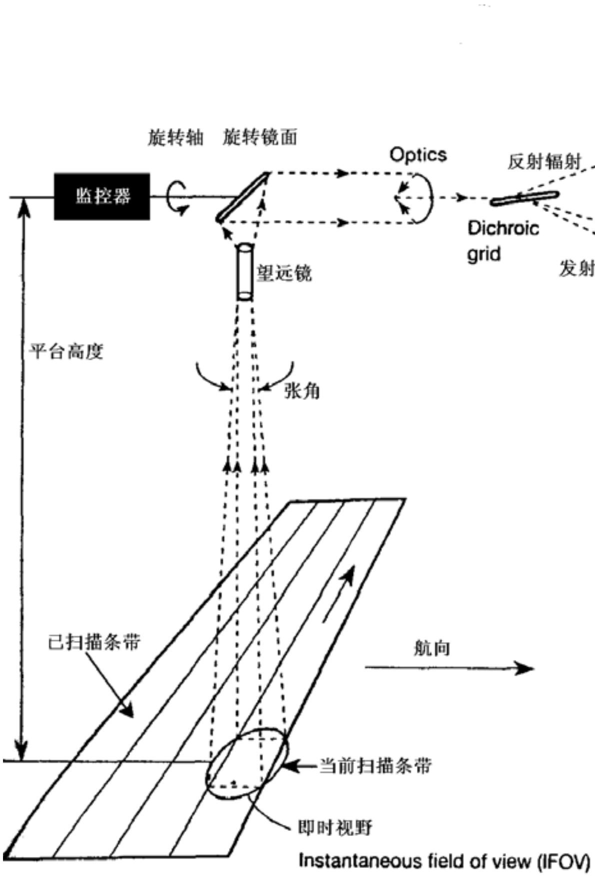


图 12-2：多光谱扫描仪示意图[Curran]

从遥感的定义中可以看出，首先，遥感器不与研究对象直接接触，也就是说，这里的“遥”并非指“遥远”；其次，遥感的目的是为了得到研究对象的特征信息；最后，通过传感器装置得到的数据，在被使用之前，还要经过一个处理过程。图 12-3 描述了从获取遥感数据到应用的过程。

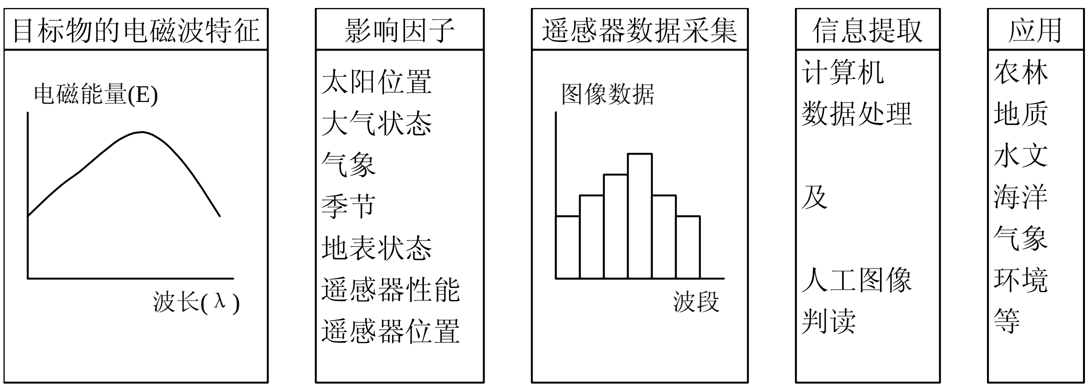


图 12-3：遥感数据过程

遥感数据的处理——通常是图像形式的遥感数据的处理，主要包括纠正（包括辐射纠正和几何纠正）、增强、变换、滤波、分类等功能，其目的主要是为了提取各种专题信息，如

土地建设情况、植被覆盖率、农作物产量和水深等等(图 12-4)。遥感图像处理可以采取光学处理和数字处理两种方式，数字图像处理由于其可重复性好、便于与 GIS 结合等特点，目前被广泛采用。下面简单介绍数字图像处理的主要功能。



图 12-4：武汉市东湖附近 TM 合成图像，长江上的轮船清晰可见

（合成方案：R=TM7，G=TM4，B=TM2，经增强处理）

1）图像纠正

图像纠正是消除图像畸变的过程，包括辐射纠正和几何纠正。辐射畸变通常由于太阳位置，大气的吸收、散射引起；而几何畸变（图 12-5）的原因则包括遥感平台的速度、姿态变化，传感器，地形起伏等，几何纠正包括粗纠正和精纠正两种，前者根据有关参数进行纠正；而后者通过采集地面控制点（GCPs, Ground Control Points），建立纠正多项式，进行纠正。

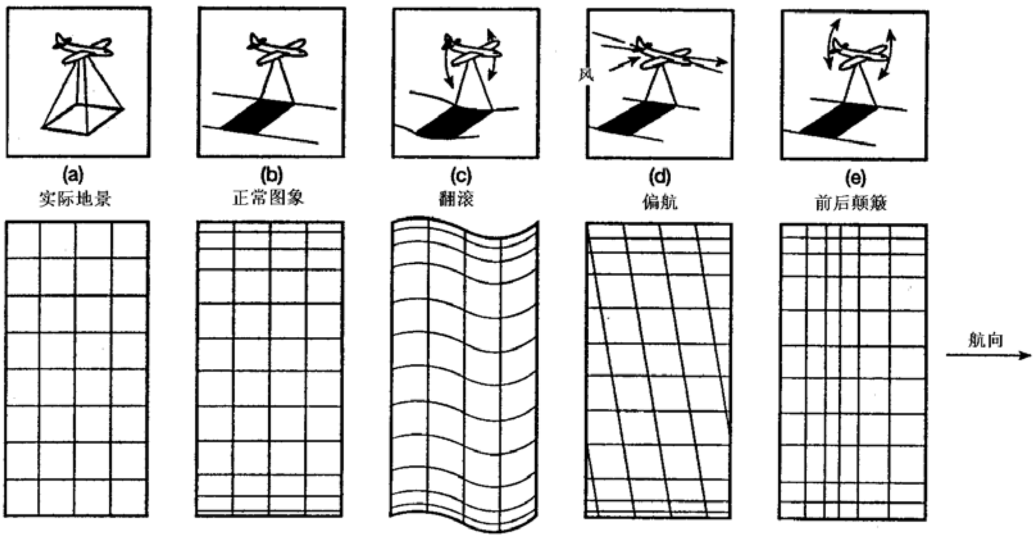


图 12-5：遥感图像几何畸变的各种情形[Lillesand and Kiefer]

2）增强

增强的目的是为了改善图像的视觉效果，并没有增加信息量，包括亮度、对比度变化以及直方图变换等。

3）滤波

滤波分为低通滤波、高通滤波和带通滤波等，低通滤波可以去除图像中的噪声，而高通滤波则用于提取一些线性信息，如道路，区域边界等。滤波可以在空域上采用滤波模板操作，

也可以在频域中进行直接运算。

4）变换

包括主成分分析(Principal Component Analyst)，色度变换以及傅立叶变换等，还包括一些针对遥感图像的特定变换，如缨帽变换。

5）分类

利用遥感图像的主要目的是为了提取各种信息，一些特定的变换可以用于提取信息，但是最主要的手段则是通过遥感图像分类(Classification)。计算机分类的基本原理是计算图像上每个象元的灰度特征，根据不同的准则，进行分类。遥感图像分类有两类方法，即监督分类(Supervised Classification)和非监督分类(Unsupervised Classification)，前者需要事先确定各个类别及其训练区(Training Area)，并计算训练区象元灰度统计特征，然后将其它象元归并到不同类别；后者则直接根据象元灰度特征之间的相似和相异程度进行合并和区分，形成不同的类别。典型的监督分类算法有最小距离法、最大似然法、平行六面体法等；而 K-均值聚类属于非监督分类；将人工神经网络（ANN, Artificial Neural Network）应用于遥感分类，在有些情况下，可以达到较好的分类效果。

遥感的出现，扩展了人类对于其生存环境的认识能力，较之于传统的野外测量和野外观测得到的数据，遥感技术具有以下优点：

- 增大了观测范围；
- 能够提供大范围的瞬间静态图像，用于监测动态变化的现象；
- 能够进行大面积重复观测，即使是人类难以到达的偏远地区；
- 大大“加宽”了的人眼所能观察的光谱范围，遥感使用的电磁波波段从 X 光到微波，远远超出了可见光范围；而雷达遥感由于使用微波，可以不受制于昼夜、天气变化，进行全天候的观测；
- 空间详细程度高，航空相片的空间分辨率可以高达厘米级甚至毫米级。

与航空遥感相比，航天遥感能够进行连续的、全天候的工作，提供更大范围的数据，其成本更低，是获取遥感数据的主要方式；而航空遥感主要应用于临时性的、紧急的观测任务以获得高精度数据。目前，世界上许多国家都已经发射了服务于不同目的的各种遥感卫星，其遥感器的空间分辨率和光谱分辨率也都各异，形成了从粗到细的对地观测数据源系列，可以用于监测从土地利用、农作物生长、植被覆盖到洪水、森林火灾、污染等现象的信息以及其动态变化。

总之，利用遥感技术，可以更加迅速、更加客观地监测环境信息；同时，由于遥感数据的空间分布特性，可以作为地理信息系统的一个重要的数据源，以实时更新空间数据库。

表 12-1：几种常用的遥感卫星及其遥感器参数

卫星传感器	波段(μm)	空间分辨率	覆盖范围	周期	主要用途
Landsat TM	0.45-0.52 0.52-0.60 0.63-0.69 0.76-0.90 1.55-1.75 10.4-12.4 2.05-2.35	30m(1-5,7 波段)	185km×185km	16 天	水深、水色 水色、植被 叶绿素、居住区 植物长势 土壤和植物水分 云及地表温度 岩石类型
SPOT-HRV	0.50-0.59 0.61-0.68 0.79-0.89 0.51-0.73	20m 20m 20m 10m	60km×60km	26 天	水色、植物状况 叶绿素、居住区 植物长势 制图
NOAA-VHRR	0.58-0.68 0.72-1.10	1.1km	2400km×	0.5 天	植物、云、冰雪 植物、水陆分界

	3.55-3.93 10.3-11.3 11.5-12.5		2400km		热点、夜间云 云及地表温度 大气及地表温度
IKONOS*	0.45-0.9 0.45-0.52 0.52-0.60 0.63-0.69 0.76-0.90	0.82m 4m 4m 4m 4m	11km× 11km	14 天	

2．全球定位系统简介

全球定位系统(GPS, Global Positioning System)是利用人造地球卫星进行点位测量导航技术的一种，其它的卫星定位导航系统有俄罗斯的 GLONASS，欧洲空间局的 NAVSAT，国际移动卫星组织的 INMARSAT 等等。GPS 全称是 NAVSTAR(NAVigation Satellite Timing And Ranging)/GPS，由美国军方组织研制建立，从 1973 年开始实施，到九十年代初完成。

2．1 GPS 系统介绍

GPS 系统包括三大部分：空间部分——GPS 卫星星座；地面控制部分——地面监控系统；用户设备部分——GPS 信号接收机。

2．1．1 GPS 卫星及其星座

GPS 由 21 颗工作卫星和 3 颗备用卫星组成，它们均匀分布在六个相互夹角为 60 度的轨道平面内，即每个轨道上有四颗卫星。卫星高度离地面约 20000 公里，绕地球运行一周的时间是 12 恒星时，即一天绕地球两周。GPS 卫星用 L 波段两种频率的无线电波（1575.42MHz 和 1227.6MHz）向用户发射导航定位信号，同时接收地面发送的导航电文以及调度命令。

2．1．2 地面控制系统

对于导航定位而言，GPS 卫星是一动态已知点，而卫星的位置是依据卫星发射的星历——描述卫星运动及其轨道的参数——计算得到的。每颗 GPS 卫星播发的星历是由地面监控系统提供的，同时卫星设备的工作监测以及卫星轨道的控制，都由地面控制系统完成。

GPS 卫星的地面控制站系统包括位于美国科罗拉多的主控站以及分布全球的三个注入站和五个监测站组成，实现对 GPS 卫星运行的监控。

2．1．3 GPS 信号接收机

GPS 信号接收机（图 12-6）的任务是，捕获 GPS 卫星发射的信号，并进行处理，根据信号到达接收机的时间，确定接收机到卫星的距离。如果计算出四颗或者更多卫星到接收机

---

* IKONOS卫星于 1999 年发射成功，其提供的高分辨率遥感数据具有良好的应用前景。相关信息来自 http://www.spaceimage.com。

的距离，再参照卫星的位置，就可以确定出接收机在三维空间中的位置。



图 12-6：GARMIN 手持式 GPS 接收机

2．2 GPS 定位基本原理

GPS 定位基本原理是利用测距交会确定点位。如图 12-7 所示，一颗卫星信号传播到接收机的时间只能决定该卫星到接收机的距离，但并不能确定接收机相对于卫星的方向，在三维空间中，GPS 接收机的可能位置构成一个球面；当测到两颗卫星的距离时，接收机的可能位置被确定于两个球面相交构成的圆上；当得到第三颗卫星的距离后，球面与圆相交得到两个可能的点；第四颗卫星用于确定接收机的准确位置。因此，如果接收机能够得到四颗 GPS 卫星的信号，就可以进行定位；当接收到信号的卫星数目多于四个时，可以优选四颗卫星计算位置。

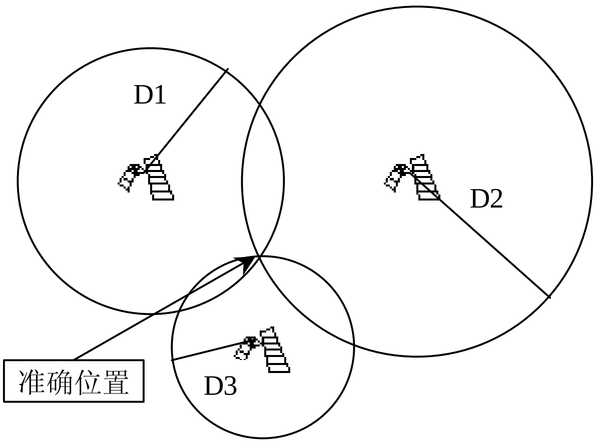


图 12-7：测距交会定位示意图

2．3 GPS 误差和纠正

造成 GPS 定位误差的因素有很多，如由于卫星轨道变化以及卫星电子钟不准确以及定

位信号穿越电离层和地表对流层时速度的变化等引起的误差，但是 GPS 定位中最为严重的误差则是由于美国军方人为降低信号质量造成的，这种误差可高达 100 米。

美国为了防止未经许可的用户把GPS用于军事目的，实施了各种技术。首先GPS卫星发射的无线电信号包括两种不同的测距码，即P码（也称精码）和C/A码（也称粗码），相应两种测距码GPS提供两种定位服务方式，即精密定位服务(PPS)和标准定位服务(SPS)，前者的服务对象主要是美国军事部门和其它特许部门，后者则服务于广大民间用户。此外，通过使用SA（Selective Availability，选择可用性）技术，C/A码的定位精度从 20 米降低至 100 米*；而AS(Anti-spoofing，反电子欺骗)技术用于对P码进行加密，当实施AS时，非特许用户不能得到P码。

上述的人为误差给 GPS 的民用造成了障碍，但是可以通过差分纠正来消除。差分纠正是通过两个或者更多的 GPS 接收机完成的，其方法是在某一已知位置，安置一台接收机作为基准站接收卫星信号，然后在其它位置用另一台接收机接收信号，由前者可以确定卫星信号中包含的人为干扰信号，而在后者接收到的信号中减去这些干扰，即可以大大降低 GPS 的定位误差。

3．GIS 与遥感的集成及具体技术

简而言之，地理信息系统是用于分析和显示空间数据的系统，而遥感影象是空间数据的一种形式，类似于 GIS 中的栅格数据。因而，很容易在数据层次上实现地理信息系统与遥感的集成，但是实际上，遥感图像的处理和 GIS 中栅格数据的分析具有较大的差异，遥感图像处理的目的是为了提取各种专题信息，其中的一些处理功能，如图像增强、滤波、分类以及一些特定的变换处理（如陆地卫星影象的 KT 变换）等，并不适用于 GIS 中的栅格空间分析，目前大多数 GIS 软件也没有提供完善的遥感数据处理功能，而遥感图像处理软件又不能很好地处理 GIS 数据，这需要实现集成的 GIS。

在软件实现上，GIS 与遥感的集成，可以有以下三个不同的层次[Ehlers]：

- 分离的数据库，通过文件转换工具在不同系统之间传输文件；
- 两个软件模块具有一致的用户界面和同步的显示；
- 集成的最高目的是实现单一的、提供了图像处理功能的 GIS 软件系统。

在一个遥感和地理信息系统的集成系统中，遥感数据是 GIS 的重要信息来源，而 GIS 则可以作为遥感图像解译的强有力的辅助工具，具体而言，有以下的应用方面[J. C. Hinton]：

1） GIS 作为图像处理工具

将 GIS 作为遥感图像的处理工具，可以在以下几个方面增强标准的图像处理功能：

1．1）几何纠正和辐射纠正

在遥感图像的实际应用中，需要首先将其转换到某个地理坐标系下，即进行几何纠正。通常几何纠正的方法是利用采集地面控制点建立多项式拟合公式，它们可以从 GIS 的矢量数据库中抽取出来，然后确定每个点在图像上对应的坐标，并建立纠正公式。在纠正完成后，可以将矢量点叠加在图像上，以判断纠正的效果。为了完成上述功能，需要系统能够综合处理栅格和矢量数据。

一些遥感影象，会因为地形的影响而产生几何畸变，如侧视雷达（Dideways-looking radar）图像的叠掩（Layover）、阴影（Shadow）、前向压缩（Foreshortening）等等，进行纠正、解译时需要使用 DEM 数据以消除畸变。此外，由于地形起伏引起光照的变化，也会在遥感图像上表现出来，如阴坡和阳坡的亮度差别，可以利用 DEM 进行辐射纠正，提高图像

---

* 美国已于 2000 年 5 月取消SA政策，使得单点定位精度可以达到 20-30 米。

分类的精度。

1．2）图像分类

对于遥感图像分类，与 GIS 集成最明显的好处是训练区的选择，通过矢量/栅格的综合查询，可以计算多边形区域的图像统计特征，评判分类效果，进而改善分类方法。

此外，在图像分类中，可以将矢量数据栅格化，并作为“遥感影象”参与分类，可以提高分类精度，例如，考虑到植被的垂直分带特性，在进行山区的植被分类时，可以结合 DEM，将其作为一个分类变量。

1． 3）感兴趣区域的选取

在一些遥感图像处理中，常常需要只对某一区域进行运算，以提取某些特征，这需要栅格数据和矢量数据之间的相交运算。

2） 遥感数据作为 GIS 的信息来源

数据是 GIS 中最为重要的成分，而遥感提供了廉价的、准确的、实时的数据，目前如何从遥感数据中自动获取地理信息依然是一个重要的研究课题，包括：

2．1）线以及其它地物要素的提取

在图像处理中，有许多边缘检测（Edge Detection）滤波算子，可以用于提取区域的边界（如水陆边界）以及线形地物（如道路、断层等），其结果可以用于更新现有的 GIS 数据库，该过程类似于扫描图像的矢量化。

2．2）DEM 数据的生成

利用航空立体像对(Stereo Images)以及雷达影象，可以生成较高精度的 DEM 数据。

2．3）土地利用变化以及地图更新

利用遥感数据更新空间数据库，最直接的方式就是将纠正后遥感图像作为背景底图，并根据其进行矢量数据的编辑修改。而对遥感图像数据进行分类，得到的结果可以添加到 GIS 数据库中。因为图像分类结果是栅格数据，所以通常要进行栅格转矢量运算；如果不进行转换，可以直接利用栅格数据进行进一步的分析，则需要系统提供栅格/矢量相交检索功能。

因为遥感图像可以视为一种特殊的栅格数据，所以不难实现遥感和 GIS 的集成的工具软件——关键是提供非常方便的栅格/矢量数据相互操作和相互转换功能，但是要注意的是，由于各种因素的影响，使得从遥感数据中提取的信息不是绝对准确的，在通常的土地利用分类中，90%的分类精度就是相当可观的结果，因而需要野外实际的考察验证——在这个过程中可以使用 GPS 进行定位。此外，还要考虑尺度问题，即遥感影象空间分辨率和 GIS 数据比例尺的对应关系，例如在实践中，一个常见的问题是：地面分辨率为 30 米的 TM 数据，进行几何纠正时，需要多大比例尺的地形图以采集地面控制点坐标，而其分类结果可以用来更新多大比例尺的土地利用数据，根据经验，合适的比例尺为 1：5 万到 1：10 万，太大则遥感数据精度不够，过小则是对遥感数据的“浪费”。

4．GIS 与全球定位系统的集成及具体技术

作为实时提供空间定位数据的技术，GPS 可以与地理信息系统进行集成，以实现不同的具体应用目标：

1）定位

主要在诸如旅游、探险等需要室外动态定位信息的活动中使用。如果不与 GIS 集成，利用 GPS 接收机和纸质地形图，也可以实现空间定位；但是通过将 GPS 接收机连接在安装 GIS 软件和该地区空间数据的便携式计算机上，可以方便地显示 GPS 接收机所在位置并实时显示其运动轨迹，进而可以利用 GIS 提供的空间检索功能，得到定位点周围的信息，从而实现决策支持。

2）测量

主要应用于土地管理、城市规划等领域，利用 GPS 和 GIS 的集成，可以测量区域的面积或者路径的长度。该过程类似于利用数字化仪进行数据录入，需要跟踪多边形边界或路径，采集抽样后的顶点坐标，并将坐标数据通过 GIS 记录，然后计算相关的面积或长度数据。

在进行 GPS 测量时，要注意以下一些问题，首先，要确定 GPS 的定位精度是否满足测量的精度要求，如对宅基地的测量，精度需要达到厘米级，而要在野外测量一个较大区域的面积，米级甚至几十米级的精度就可以满足要求；其次，对不规则区域或者路径的测量，需要确定采样原则，采样点选取的不同，会影响到最后的测量结果。

3）监控导航

用于车辆、船只的动态监控，在接收到车辆、船只发回的位置数据后，监控中心可以确定车船的运行轨迹，进而利用 GIS 空间分析工具，判断其运行是否正常，如是否偏离预定的路线，速度是否异常（静止）等等，在出现异常时，监控中心可以提出相应的处理措施，其中包括向车船发布导航指令。

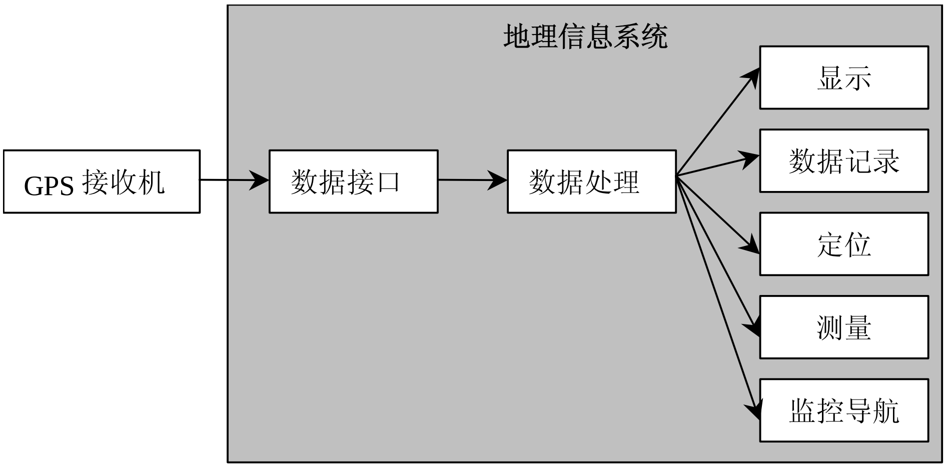


图 12-8：GIS 与 GPS 集成的系统结构模型

图 12-8 描述了 GIS 与 GPS 集成的系统结构模型，为了实现与 GPS 的集成，GIS 系统必须能够接收 GPS 接收机发送的 GPS 数据（一般是通过串口通信），然后对数据进行处理，如通过投影变换将经纬度坐标转换为 GIS 数据所采用的参照系中的坐标，最后进行各种分析运算，其中坐标数据的动态显示以及数据存储是其基本功能。

5．3S 集成综述

3S技术为科学研究、政府管理、社会生产提供了新一代的观测手段、描述语言和思维工具。3S的结合应用，取长补短，是一个自然的发展趋势，三者之间的相互作用形成了“一个大脑，两只眼睛”的框架，即RS和GPS向GIS提供或更新区域信息以及空间定位，GIS进行相应的空间分析（图12-9），以从RS和GPS提供的浩如烟海的数据中提取有用信息，并进行综合集成，使之成为决策的科学依据。

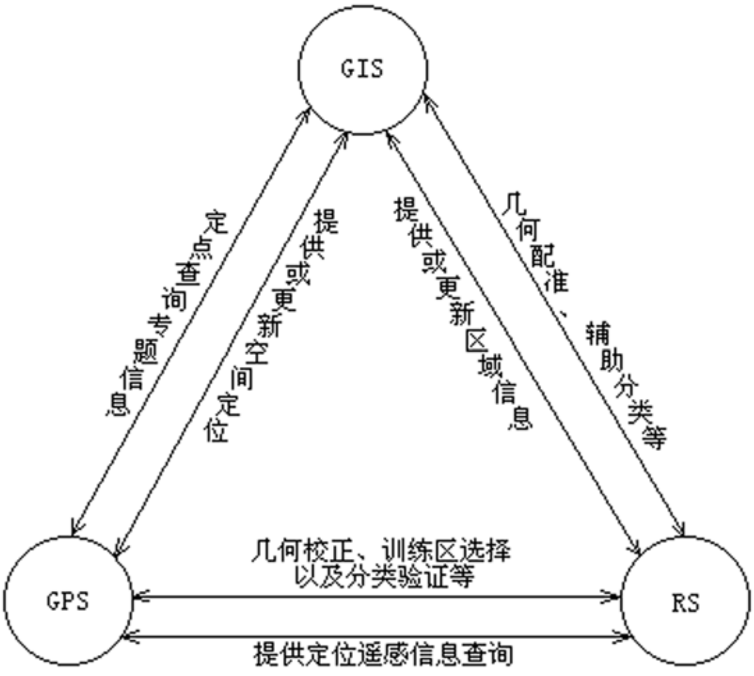


图12-9：3S的相互作用与集成[邬伦]

GIS、RS和GPS三者集成利用，构成为整体的、实时的和动态的对地观测、分析和应用的运行系统，提高了GIS的应用效率。在实际的应用中，较为常见的是3S两两之间的集成，如GIS/RS集成，GIS/GPS集成或者RS/GPS集成等，但是同时集成并使用3S技术的应用实例则较少。美国Ohio大学与公路管理部门合作研制的测绘车是一个典型的3S集成应用，它将GPS接收机结合一台立体视觉系统载于车上，在公路上行驶以取得公路以及两旁的环境数据并立即自动整理存储于GIS数据库中。测绘车上安装的立体视觉系统包括有两个CCD摄象机，在行进时，每秒曝光一次，获取并存储一对影象，并作实时自动处理。

RS、GIS、GPS 集成的方式可以在不同的技术水平上实现，最简单的办法是三种系统分开而由用户综合使用，进一步是三者有共同的界面，做到表面上无缝的集成，数据传输则在内部通过特征码相结合，最好的办法是整体的集成，成为统一的系统。

单纯从软件实现的角度来看，开发 3S 集成的系统在技术上并没有多大的障碍。目前一般工具软件的实现技术方案是：通过支持栅格数据类型及相关的处理分析操作以实现与遥感的集成，而通过增加一个动态矢量图层以与 GPS 集成。对于 3S 集成技术而言，最重要的是在应用中综合使用遥感以及全球定位系统，利用其实时、准确获取数据的能力，降低应用成本或者实现一些新的应用。

3S集成技术的发展，形成了综合的、完整的对地观测系统，提高了人类认识地球的能力；相应地，它拓展了传统测绘科学的研究领域。作为地理学的一个分支学科，Geomatics*产生并对包括遥感、全球定位系统在内的现代测绘技术的综合应用进行探讨和研究。同时，它也推动了其它一些相联系的学科的发展，如地球信息科学、地理信息科学等，它们成为“数字地球”这一概念提出的理论基础。

---

* Geomatics以及其它学科的介绍见“地球信息科学与数字地球”一章。

网络就是计算机。

Sun Microsystem Corp.

第十三章 网络地理信息系统

导读：本章首先介绍了计算机网络的基本知识，包括定义、拓扑结构、分类、网络协议等，然后介绍了因特网以及目前流行的 WWW 和 HTML 语言。
第二节讲述了基于网络的分布计算以及地理信息系统在分布计算环境下的具体实现方案。
第三节主要介绍了在 Web 上发布空间数据，即实现 WebGIS 的几种技术方案，并进行了比较。

1．计算机网络技术

1．1 网络的基本概念

1．1．1 定义

计算机网络（Computer Network）是指实现计算机之间通讯的软件和硬件系统的统称，从广义上讲，利用磁盘在两台微机之间拷贝数据也可以认为是一种特殊的网络。它的更加具体的定义是“以共享资源为目的，通过数据通讯线路将多台计算机互联而组成的系统”，共享的资源包括计算机网络中的硬件设备、软件或者数据。

1．1．2 网络拓扑结构和分类

计算机网络中，网络结点(Node)计算机或者终端的连接方式称为网络的拓扑结构，它影响着网络的设计、功能、可靠性以及通讯费用等方面，常见的拓扑结构有以下类型：（图 13-1）

- 星形拓扑结构；
- 总线拓扑结构；
- 环形拓扑结构；
- 树形拓扑；
- 完全连接结构；
- 相交环形；
- 不规则拓扑结构等等。

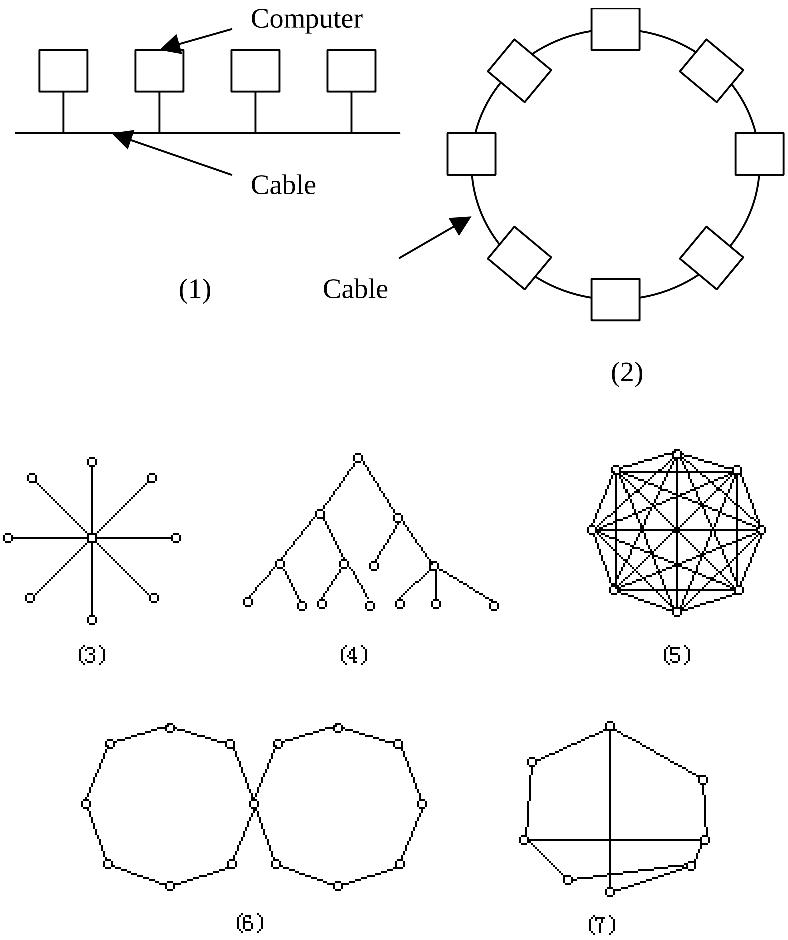


图 13-1：网络拓扑结构(A. S. Tanenbaum)
（1）总线结构，（2）环形结构，（3）星形结构，（4）树形结构，
（5）完全连接，（6）相交环形，（7）不规则拓扑结构

计算机网络的种类繁多，性能各异，按照不同的原则可以有不同的分类方式，最常见的是根据其空间分布范围的大小，区分为局域网(LAN-Local Area Network)、城域网(MAN-Metropolitan Area Network)和广域网(WAN-Wide Area Network)。

1）局域网

是指在比较小的区域——如一座办公大楼，一个校园，一个公司等内建立的计算机网络，其通讯距离较短，传输速率较快，误码率低。对于局域网而言，主要采用的拓扑结构是总线结构和环形结构，前者如IEEE 802.3，通常称为Ethernet™(以太网)，后者如IEEE 802.5，即IBM的令牌环网(Token Ring)。

2）城域网

基本上是一个大的局域网，采用类似的拓扑结构，它可以覆盖一个城市，实现信息共享。

3）广域网

广域网最根本的特点是其机器分布范围广，可以实现大范围的信息共享，笼统地讲，因特网(Internet)就属于广域网范畴。广域网可以采用多种拓扑结构，如星形，环形，树型，完全连接，相交环形，甚至是不规则的拓扑结构。

上述的网络类型，结点之间都是通过电缆进行连接，近年来无线网(Wireless Network)得到较快的发展，利用无线网，可以实现移动计算。

1．2 因特网(Internet)

因特网是指全球最大的、开放的、由众多网络相互连接而成的计算机网络，其形成和发展无疑是信息技术发展过程中最为激动人心的事情，目前大专院校，大学，政府部门，图书馆，企业以及个人都已经连接到因特网，它允许网络使用者在任何时间、任何地点查阅任何网络上的信息，它给信息传播带来了巨大便利，它正在改变着整个人类社会的生活方式。

1．2．1 因特网发展历史

因特网的前身是 1965 年开始由美国国防部设计的 ARPA(Advanced Research Project Agency)net。经过周密的策划，1969 年，分布于美国四所大学（UCLA,UCSB,SRI,UTAH）的计算机开始进行互相数据交换，标志着 ARPAnet 的诞生，后来，特别是 1983 年 TCP/IP 协议被正式应用于 ARPAnet 之后，其接入用户飞速增长。

由于 ARPAnet 的军方背景，给更加便捷的连入和访问造成不便，于是 NSF(National Science Foundation)建立了 NSFnet，其管理和速度都优越于 ARPAnet，到 1990 年，很多用户从 ARPAnet 转到 NSFnet，后来 NSFnet 与 ARPAnet 互连，其用户成指数增长。许多地区性（包括加拿大、欧洲、大洋州等）的网络联入后，标志着 Internet 的最终形成。

1．2．2 因特网提供的应用服务

在传统上，因特网主要有以下四种应用：

- 电子邮件(Electronic Mail)；
- 新闻组(News)；
- 远程登录；
- 文件传输。

直到八十年代初期，因特网还只是局限于学院、政府等等，但是 WWW(World Wide Web)的出现改变了这一切，WWW 使得一个站点可以建立一些包括文本、图片、声音甚至录象的页面，这些页面用超文本标志语言(HTML-Hyper Text Markup Language)写成，内嵌指向其它页面的链接。WWW 为用户查询、检索、浏览在因特网上发布的各种信息提供了极大的便利，为因特网带来了大量的非学院用户。实际上因特网的迅猛发展，与 WWW 应用是密不可分的，以至于今天提到 Internet 时常常就是特指 WWW 服务。后面将对 WWW 以及相关技术进行更为详细的说明。

其它一些因特网应用包括 Gopher，WAIS 等等。

1．2．3 网络协议(Protocol)和 TCP/IP

在计算机网络中，为了使得相连的计算机或终端之间能够正确的传输信息，必须有一整套关于信息传输顺序、信息格式以及内容的约定，称为计算机网络通讯协议。

在网络协议中，由国际标准化组织(ISO)发展和制定开放系统互连参考模型（OSI，Open System Interconnect Reference Model）制订了数据通讯协议的标准，尽管 OSI 模型并没有真正地在实际中被应用，但是其提供的概念和词汇被计算机网络界广泛地使用和认可。OSI 参考模型包括七个协议层来定义数据通讯的协议功能，每一层是相对独立的，完成数据传输

过程中的部分功能。

目前，由于因特网的发展，TCP/IP 作为其基础协议，应用越来越广泛。TCP/IP 狭义特指两个协议及传输控制协议(TCP，Transformation Control Protocol)和网际协议(IP，Internet Protocol)，广义指由多个与因特网相关的协议组成的 TCP/IP 参考模型，或者称为 TCP/IP 协议栈(Protocol Stack)，TCP/IP 模型类似于 OSI 参考模型，但是较为简单，没有表示层和会话层。TCP 及 IP 是其中重要的两个协议，IP 是网络层协议，而 TCP 位于传输层。

TCP/IP 应用层包括所有高层应用协议，如虚拟终端（TELNET）、文件传输(FTP，File Transfer Protocol)、电子邮件(SMTP，Simple Mail Transfer Protocol 和 POP3-PostOffice Protocol)、域名服务(DNS，Domain Name Service)以及 Web 上用来传递页面的超文本传输协议(HTTP，Hyper Text Transfer Protocol)。

OSI 参考模型结构及其和 TCP/IP 参考模型的对应关系如图 13-2：

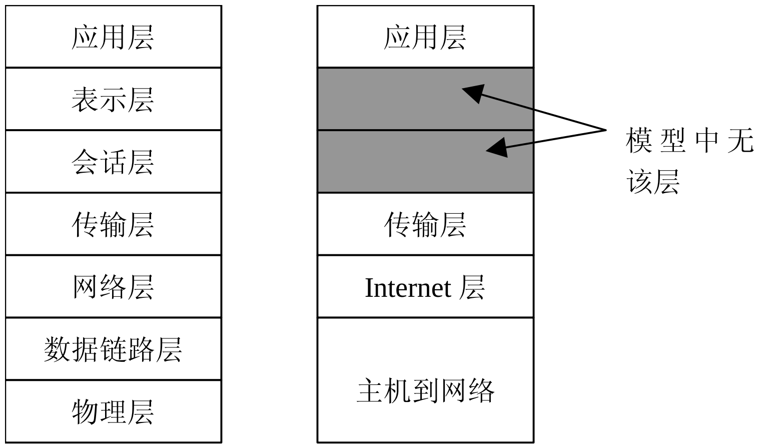


图 13-2：OSI 参考模型和 TCP/IP 参考模型

1．2．4 因特网的地址、域名

为了能够在因特网上传输信息，必须有一种方法能够识别计算机，其方案是为因特网上的每一台计算机都赋予了一个独一无二的 32 位地址，称为 IP 地址，该 32 位地址通采用用点号分割的十进制数表示，每个字节为一个十进制数值，其数值范围从 0 到 255，这样，IP 地址就表现为如下的形式：202.112.7.12（北京大学 WWW 服务器 IP 地址）。在因特网上，计算机之间的通讯是通过 IP 地址来进行的。由于因特网的迅速发展，有越来越多的计算机连入，现有的 32 位 IP 地址已经不太能够满足其进一步发展的需要，因此，ITEF(Internet Engineering Task Force, 因特网工程任务组)开始制订更新的 IP 协议标准，称为 Ipv6，根据该标准，IP 地址将采用 16 个字节来标识。

虽然计算机之间的通讯是利用其各自的IP地址来实现的，但是人们在使用过程中，不愿意总是使用难以记忆的数字，而是希望使用更为方便的文字名称。这样网络上的各个主机除了IP地址之外，通常还有一个名称来表示。该名称称为域名（Domail Name），与IP地址一样，也是唯一的。这种在因特网上使用的名称系统称为域名系统（DNS，Domain Name System）。DNS是一个层次状的命名系统，通过一个数据库对连接因特网的所有主机进行名称和地址的管理。域名是以点号分割的字符串，如PKU.EDU.CN和OPENGIS.ORG等等。通常最右边的字串称为顶级域名，顶级域名有两类，即类顶级域名和国家顶级域名，前者有*com*(商业)，*edu*(教育界)，*gov*(政府机构)，*int*（特定的国际组织），*mil*（军队系统），*net*（网络提供者），*org*（非赢利机构）等，类顶级域名应用于美国的站点；后者用两个英文字母表示一个国家和地区，如*cn*（中国），*ca*（加拿大），*jp*（日本）等等。

1．2．5WWW、HTML 和 URL

正如前面描述的，万维网（WWW）的出现对于因特网来说具有划时代的意义，使得因特网的应用走出专业化。WWW，通常又称为 Web 或 3W，是存储在全世界的 Internet 计算机中、彼此关联、数以百万计的文档集合，它是世界上最大的电子信息仓库。在 WWW 应用中，Web 文档存放于一台或者多台计算机上，称为 Web 服务器(Web Server)，用户利用浏览器(Browser)来访问 Web 站点。

Web 的第一个版本于 1992 年 1 月在瑞士的日内瓦首先开始运行，研究人员可以从 CERN(欧洲粒子物理研究所)的站点上访问 Web 信息。Web 真正流行应该归功于 Mosaic，第一个基于图形的浏览器，该浏览器是由 NCSA(美国国家超级计算应用中心)设计编写的，其操作简便、界面友好的特点，使得人们可以方便的访问 Web 信息。目前广为流行的浏览器是 Microsoft Internet Explorer 和 Netscape Navigator。

在 Web 站点上，采用 HTML 编写文档，通常称为超文本(Hyper Text)，“超”的意思是指在文档中包含对其它文档的超链接（Hyper Link），使用户阅读时可以通过关键字跳转到其它的文档。由于允许在超文本中加入音乐以及图像等多媒体数据，又诞生了超媒体(Hyper Media)这个概念。HTML 是一种标记语言，与系统平台和应用程序无关，它通过特殊的标记和参考字符来设定 Web 文档的结构或者格式，并且有特殊的 HTML 元素（如定位元素 Anchor）来定义超链接，同时可以利用 HTML 表单元素从用户那里获取有关信息。在用户访问 Web 站点时，浏览器和 Web 服务器之间通过超文本传输协议(HTTP)来发送请求和信息。

在浏览 Web 时，首先会碰到 URL(Uniform Resource Locator,统一资源定位器)这一概念，它是 Web 文档的因特网地址，它与文档的关系就如同于书籍中目录和正文的关系。URL 的语法相当简单，由两部分组成，形如：

<方法>：<方法指定部分>（<scheme>:<scheme-specific-part>）

一个 HTTP URL 如下：

http://<host>:<port>?<search_part>

其中 host 部分由服务器的地址（IP 地址或者域名）、文件路径、文件名组成；port 部分是 HTTP 端口号，其缺省为 80，通常可以忽略；而 search_part 部分往往是浏览器向服务器发出的请求参数。下面是一个 HTTP URL 的例子：

http://www.pku.edu.cn/news1/index.htm

除了 Web 以外，FTP、Gopher 等也采用 URL 进行资源定位。

1．2．6 企业内部互连网

在九十年代后半段，随着因特网的发展，出现了企业内部互连网（Intranet）这个崭新的概念，简而言之，Intranet 是企业自己的内部网络，不过这个网络同样采用了基于 Internet 的工具，如 Email，Web 浏览器和文件传输等等，因而使得一个企业内部充分享用因特网的各种优势，如信息的共享和管理，实时通讯和协作，使用分布式的数据库等等。

2．分布式地理信息系统

随着计算机网络的发展，基于客户机/服务器体系结构，并在网络支持下的分布式系统结构已经成为地理信息系统的发展趋势。由于 GIS 的固有的特点，使得运行于网络上的分布式系统特别适合于构造较大规模的 GIS 应用，其应用表现在以下几个方面：

1）数据的分布：在地理信息系统中，主要数据是空间数据，由于数据生产和更新的要求，常常需要存放在空间上分离的计算机上。

2）应用功能的分布：GIS 的功能组成了由空间数据录入到输出的一个工作流程，不同的人员由于其关注的信息不同，需要不同的 GIS 功能服务对数据进行处理，将应用分布在网络上就可以解决该问题。

3）外设共享：外设的分布是服务分布的一种，由于许多 GIS 外设较为昂贵，如高精度平板扫描仪，喷墨绘图仪，大幅面数字化仪等，而通过分布式系统，可以实现这些设备的共享。

4）并行计算：在地理信息系统中，许多模型具有较高的时间复杂性，利用分布系统可以实现并行计算，缩短计算时间。

下面将描述分布式系统以及一些网络地理信息系统的模式和实现。

2．1 分布式系统和 C/S 模型

分布式系统(Distributed System)的定义是：一组独立计算机的集合，但是从用户的角度来看，如同于一台计算机。较之于集中式系统(Centralized System)以及独立的微机系统，分布式系统具有以下优点（表 13-1，表 13-2）：

表 13-1：分布式系统的优势--与集中式系统比较[A. S. Tanenbaum]

项目	描述
经济	多个微处理器提供了更好的性能/价格比
速度	一个分布式系统比大型机有更好的总体计算能力
固有的分布	一些应用系统需要运行于空间上分离的机器
可靠性	如果一台机器崩溃，整个系统还能够运行
增量发展	计算能力能够以很小的增量增长

表 13-2：分布式系统的优势--与微机系统比较[A. S. Tanenbaum]

项目	描述
数据共享	允许多个用户访问同一数据库
设备共享	允许多个用户共享昂贵的外设
通讯	使人与人之间的通讯变得简单，如电子邮件
灵活性	以最有效的方式将工作量分布在所有可用的机器上

在具有上述优势的同时，由于分布式系统的复杂性，对软件硬件提出了更高的要求，它也具有以下不足之处（表 13-3）：

表 13-3：分布式系统的不足之处[A. S. Tanenbaum]

项目	描述
软件	当前支持分布式系统的软件较少
网络	网络可能充满了或者引发其它的错误
安全性	容易使得需要保密的数据被访问

客户机/服务器(C/S,Client/Server)模型是一种分布式系统结构，在该体系中，客户端通常是同最终用户交互的应用软件系统，而服务器由一组协作的过程构成，为客户端提供服务。客户机和服务器通常运行相同的微内核，一个客户机/服务器机制可以有多个客户端，或者多个服务器，或者兼而有之。客户机/服务器模式基于简单的请求/应答协议，即客户端向服务器提出信息处理的请求，服务器端接收到请求并将请求解译后，根据请求的内容执行相应操作，并将操作结果传递回客户端(图 13-3)。客户机/服务器体系的优点在于简单和高效。HTTP、FTP 等协议都是遵循客户机/服务器的模式。在网络地理信息系统的实现中，客

户机/服务器模型是经常采用的一种方案。目前，将服务器分解形成数据服务器和应用服务器，形成三层结构，可以更好地区分数据访问操作和应用模型，这是经常被采用的一种结构。

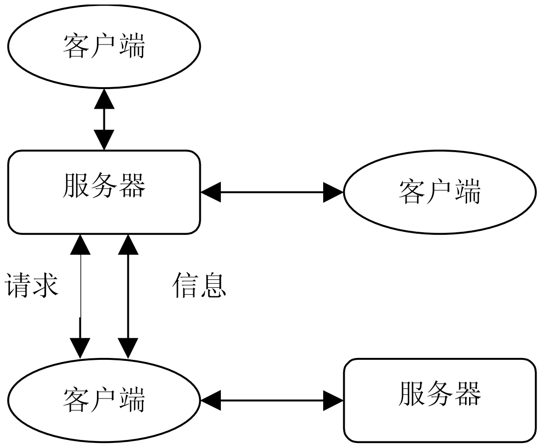


图 13-3：客户端-服务器模型

2．2 网络地理信息系统的组合方式

在分布式的网络地理信息系统中，客户机和服务器分别由相应的软件、硬件以及数据库组成，其组合可以按照数据和应用功能的分布分成五种（表 13-4）：

表 13-4：网络信息系统的组合方式[李斌]

组合方式	数据	应用功能
全集中式	中央服务器	中央服务器
数据集中式	中央服务器	客户端
功能集中式	客户端	中央服务器
全分布式	客户端	客户端
函数库分布式	客户端或者服务器	中央服务器存储，客户端动态连接执行

2．2．1 全集中式

全集中式的地理信息系统把软件、数据库管理系统和数据库全部集中在中央服务器上，客户系统只负责用户界面功能，即获得用户指令并传递给服务器，显示查询结果，提供系统的辅助功能（图 13-4）。常用的客户设置有三种，第一种是以 X-server 为代表的，只负责表现逻辑的客户系统，所有的数据处理和运算均在服务器上执行，客户端由专门的 X 终端或者 X 模拟器通过 X 协议实现用户与服务器之间的通讯；第二种是以 ArcView 为代表的客户软件系统，这类系统除了提供一般的用户界面以外，还具有相当强的分析和处理功能。ArcView 可以与 ESRI 的 ARC/INFO、与作为服务器的 SDE、Arcstorm，通过网络软件系统 NFS 或者网络 API 构成网络地理信息系统；第三种是目前在 WebGIS 上广泛采用的客户系统，用户界面功能由浏览器执行，WebGIS 在后面还将进一步的描述。

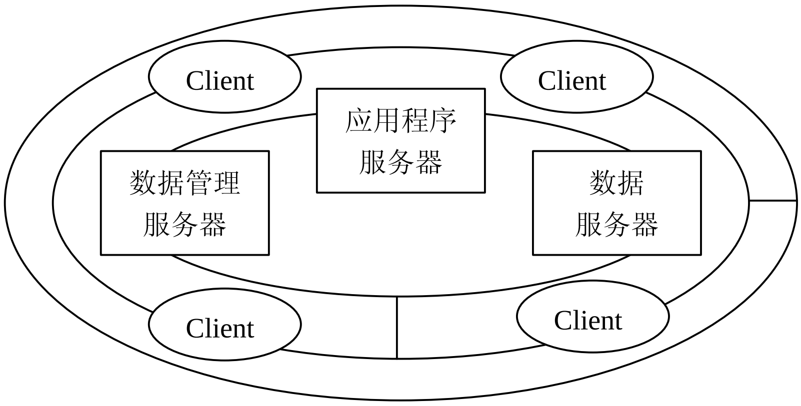


图 13-4：全集中式网络信息系统[李斌]

2．2．2 数据集中式

网络系统专门设置集中的数据存储和管理服务，网络的其它部分成为数据客户，它们一般都是带有一定功能的地理信息系统软件（图 13-5）。简单的数据服务可以由网络软件系统（如 NFS）提供，大型的管理系统则需要功能完备和高性能的数据服务器，如 Arcstorm 和 Oracle。由于目前许多数据库管理系统开始支持面向对象的数据模型，更加方便于空间数据的管理，以建立数据集中式的网络地理信息系统。

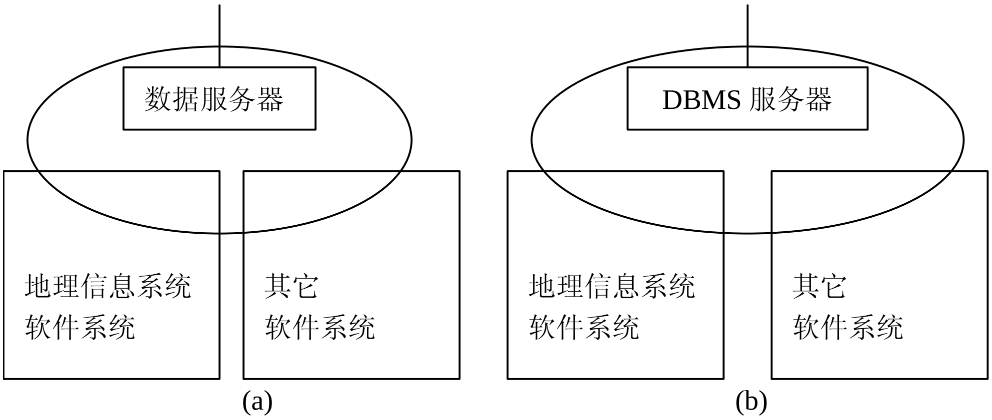


图 13-5：数据集中式地理信息系统组合模式[李斌]
(a)采用网络文件系统，(b)采用数据库服务器

2．2．3 功能集中式

与数据集中式相反，功能集中式的网络信息系统把绝大部分的功能集中在一个或者几个容量大、性能高的服务器上，由它们负责所有的分析和处理，数据则分散到客户端存储和管理（图 13-6）。由于在大多数 GIS 应用中，数据量一般比较庞大，采用这种方式，会增加网络的传输量，从而降低整个系统的性能。

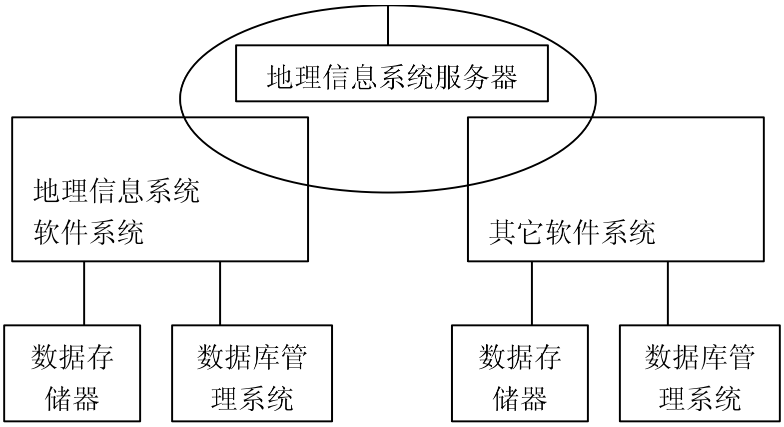


图 13-6：功能集中式的地理信息系统[李斌]

2．2．4 全分布式

全分布系统是原有的非网络化的信息系统自然进化的结果。在全分布系统中，各个子系统具有完备的数据库及地理信息系统软件和其它应用软件，在网络中同时扮演客户和服务器的角色。各个子系统的软硬件环境和特性及拥有的数据都很可能不一样，但同时又有很密切的联系和互补性。系统的集成，通过网络操作系统及各子系统提供的 API 实现。实现全分布式的网络地理信息系统，往往需要基于已有的系统平台进行二次开发，使它们能够相互协作。

2．2．5 函数库服务器

传统的软件系统一般是静态的，为了提供更多的功能，系统变得越来越大，而实际上，对于每一个用户而言，通常只是需要有限的几个功能，这样就造成了系统资源的浪费。对于集中式系统而言，系统的扩大将加大中央服务器的负担，造成系统性能下降，而全分布系统实现又较为复杂。函数服务器把优化的功能函数存储在服务器上，通过网络按用户要求动态合成应用软件，并使其在客户机上运行。从而从根本上改变了传统的资源分配和软件运行及维护方式。基于分布构件模型（CORBA 或 DCOM）构造的软件系统可以在一定程度上实现函数库服务器。

2．3 网络地理信息系统的概念设计

从逻辑上讲，基于用户的功能需求进行概念设计是大多数计算机网络软硬件配置设计的第一步。概念设计应当基于如下要点：

- 对用户的应用意图有充分的了解；
- 掌握计算机硬件和计算机通讯的基本概念知识；
- 了解已经存在的并有可能影响设计的条件，包括现有软件系统以及通讯设施；
- 对于有可能影响系统设计的计算机技术及其发展的现有状况应有充分认识。

概念设计应当对在地理信息系统网络配置中的硬件、软件的功能和作用给予说明，并对设备和应用之间的关系给予描述，应当能够在网络系统实现时，有利于系统功能重新审核以及修改，有利于系统价值的估算，并且为更详细的特殊定义提供基础，同时还应当对系统能

够在新的应用、新的用户和扩展数据库方面给予描述。

进行网络地理信息系统概念设计时，集中式、分布式和处理功能必须与特定的地址条件、用户的应用需求相适应。下面给出了一些典型概念设计以及其说明。（表 13-5）

表 13-5：一些网络地理信息系统的概念设计描述

图示	系统和组织类型	地址分布	数据传输操作	升级可能性
一	适用于地方政府、工程单位的单个部门的少量用户用于数据获取、地图生产	单一地址	与其它已有站点之间的数据传输没有特殊的要求，数据可以通过磁带等介质或者进行批量传输	可以升级为基于主机的网络或者具有其它智能工作站的网络
二	作为小型多用户系统，用来支持数据获取、地图生产、数据库和图形查询应用，服务于政府机构或企事业单位的一个或几个部门	拥有大量外围处理单元被分配在一建筑物的一层或者相邻几层，远程用户通过拨号上网与系统连接	其它来源的数据可以通过磁带等介质或者直接进行批量传输	一些附加设备可以添加在异步端口上，可以升级为局域网
三	适用于中型到大型的政府或企事业单位的组织结构，用户通过进入大型主机数据库进行查询、分析和制图	在同一建筑物内，多个地址直接相连；新增加的工作站或其它设备可以通过远程联结与图形控制器以及通讯处理器相连	周期性数据交换可以通过磁带等介质进行，也可以利用网关与其它计算机网络进行传输，微机数据的传输容易建立	设备可以灵活增加以达到大型主机处理器的计算容量
四	用于小型到中型的政府部门或公司单位	分配于同一建筑物内	可以通过微机桌面传输装置或拨号上网来进行数据交换	增设的微机和服务器可以容易的添加到环形网络上
五	适用于需要大量制图与地理分析的中型到大型的政府、公共事业和公司单位组织的多个分组部门	同一建筑物内或者临近建筑物之间的局域网上，其它用户通过远程通讯线路与局域网相连	数据传输在网络上进行，远程通过批量数据文件或磁带交换，也可通过网关连接到其它网络	局域网可以通过中继器扩展到更大的地区，多个局域网可以通过网桥连接在一起
六	设置在局域网上，支持具有主机处理单元的多用户，适用于多部门背景	同一建筑物内或者临近建筑物之间的局域网上，其它用户通过远程通讯线路与局域网相连	以主机来维持主数据库，可以进行批量文件或磁带的数据传输	可以增加外设，工作站和处理单元直至达到网络限制，可以通过中继器扩展其它建筑物，多个局域网可以通过网桥连接在一起，可以增加网关连接多种网络

七	在局域网上配置多个处理单元和外设，以支持多个功能相关组织	处理单元和外设分布在多个建筑物内，有些设备通过远程线路与局域网相连	用户可以访问局域网上的分布式数据，可以通过网关进行其它网络的批量或实时数据访问	可以增加外设，工作站和处理单元直至达到网络限制，局域网扩展可以通过中继器或网桥实现
八	一个系统网络中的多用户数据共享，或者是需要较大数据处理能力的具有地理分布的某个组织	多个处理单元通过远程线路相连	单个地址的数据更新以批量文件方式进行，多个地址之间的相互访问依靠远程线路传输	每个地址都可以扩展，增加的处理单元可以加入到远程网络中

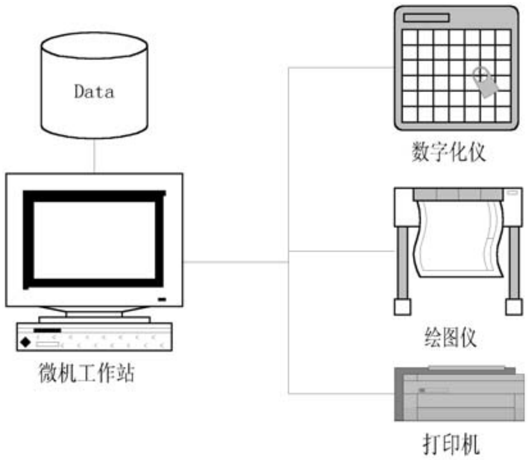


图 13-7：（一）拥有外设的超级微机工作站

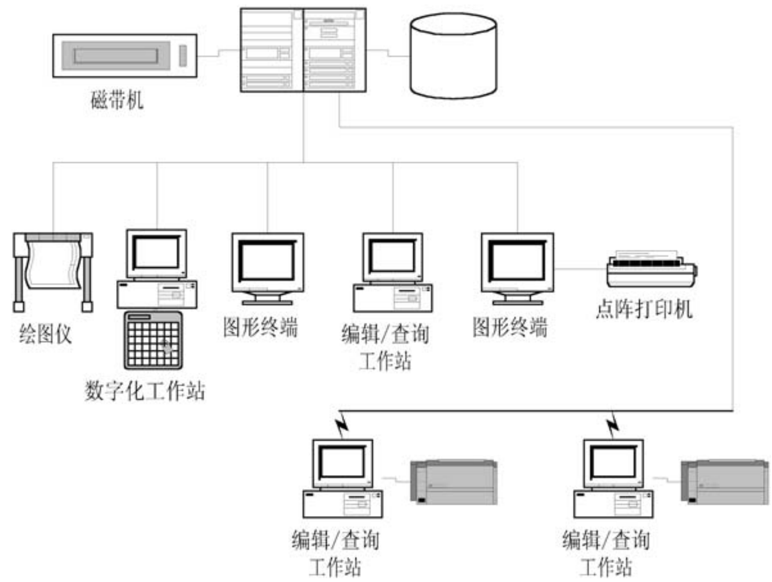


图 13-8：（二）通过异步线路支持设备的基于主机的集中式系统

网络互连设备简介

1） 中继器(Repeater)：是最简单的网间连接器，提供对信号的放大和转发，它只能连接具有相同物理协议的 LAN。中继器主要用于扩充 LAN 电缆段的距离，在同一个 LAN 中，也可以采用 LAN 延长介质长度。

2） 网桥(Bridge)：网桥是在数据链路层实现 LAN 互连的存储转发设备，它独立于高层协议，可以实现异构型局域网的互连。

3） 路由器(Router)：路由器工作在 OSI 模型的网络层，能够获得更多的网络信息，为来到的信息找到“最佳路由”。路由器结构比网桥复杂，速度也慢，但是具有更大的灵活性和更强的异种网络互连能力。

4） 网关(Gateway)：又称为协议转换器，工作在 OSI 模型的传输层及以上层，用于互连不同体系结构的网络或者媒体。它不仅要连接分离的网络，还必须确保从一个网络传输的数据与另一个兼容。

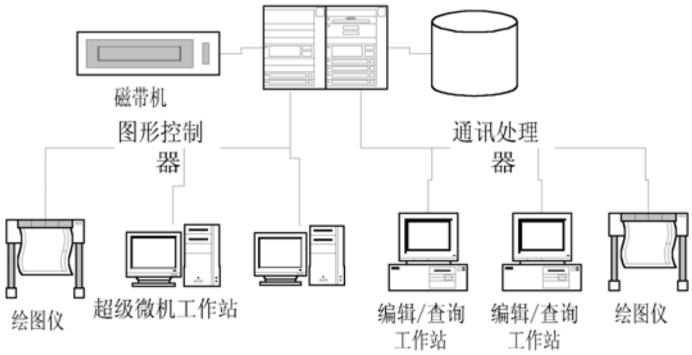


图 13-9：（三）通过通讯控制器支持外围设备的基于大型主机的系统

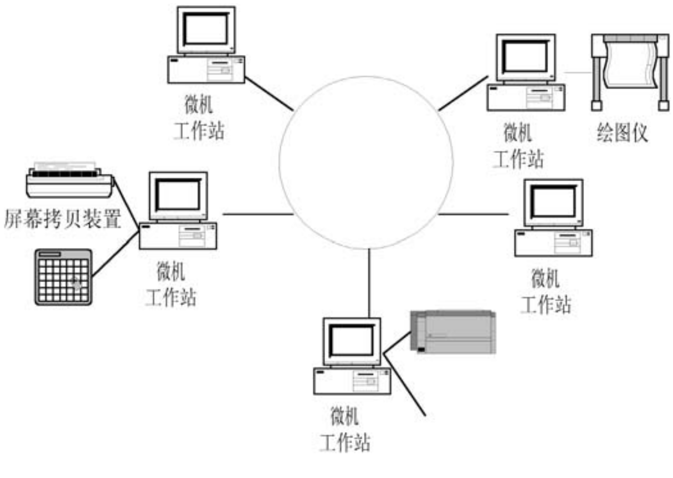


图 13-10：（四）令牌环网络微机系统

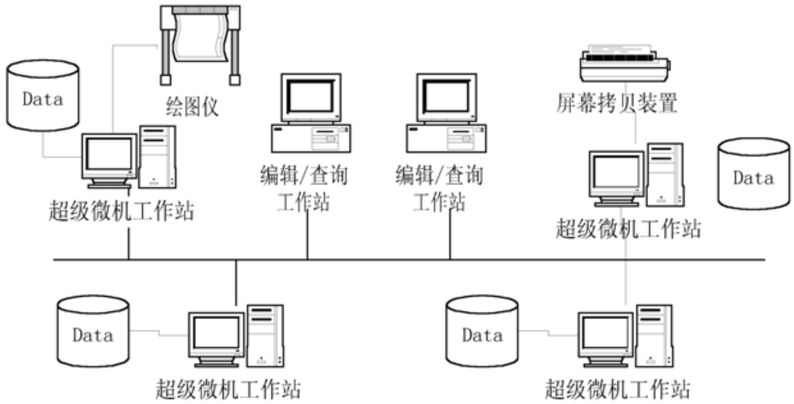


图 13-11：（五）局域网上具有超级微机工作站的分布式处理和数据库系统

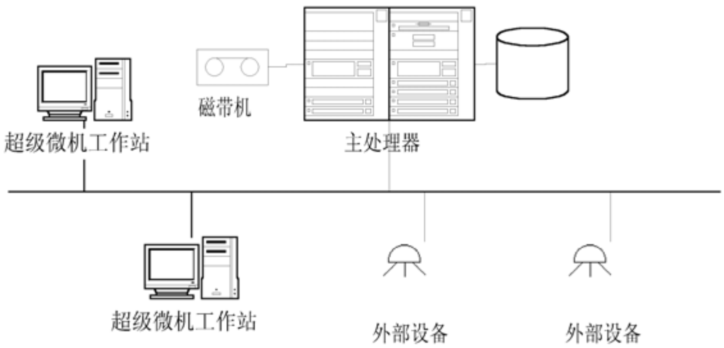


图 13-12：（六）具有主处理器，超级微机工作站和外设的局域网

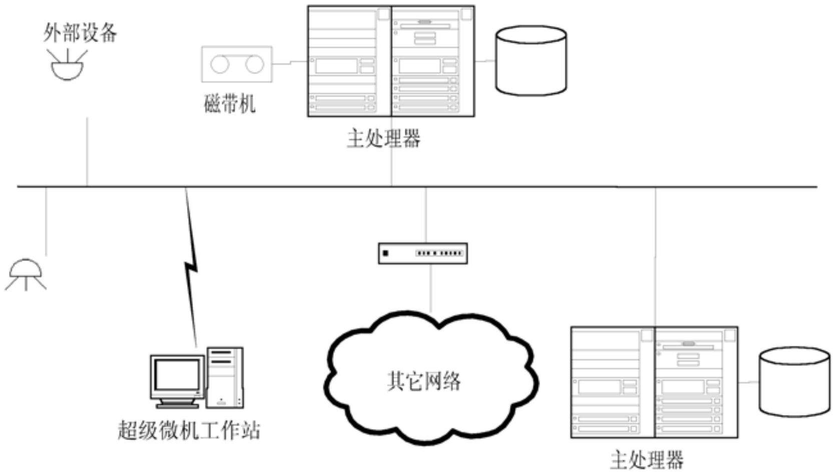


图 13-13：（七）具有多处理单元和中央数据存储的局域网

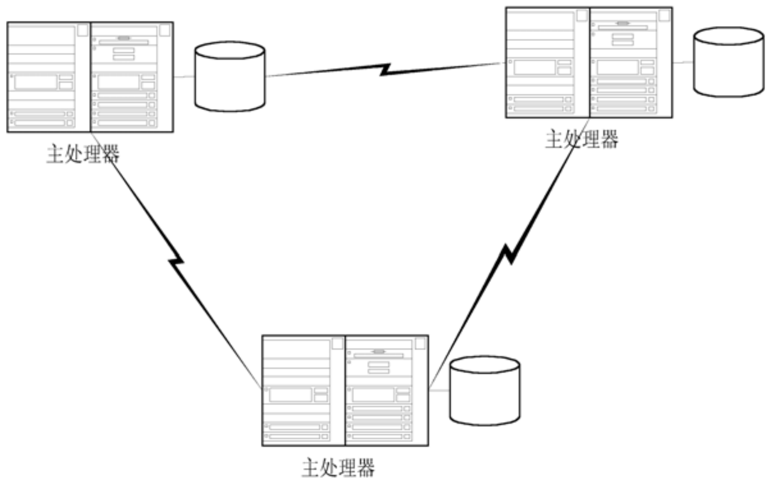


图 13-14：（八）具有分布式数据存储的远程处理

几种局域网技术标准：以太网，令牌环网和 FDDI

IEEE(电气和电子工程师协会)确定了一些局域网的标准，统称为 IEEE 802，包括 CSMA/CD(带有冲突检测的载波侦听多路存取)，令牌总线，令牌环，而 FDDI（Fiber Distributed Data Interface，光纤分布式数据接口）应用于高速 LAN。下面对几个常用的技术标准作以简单介绍。

1） IEEE 802.3：CSMA/CD，采用 CSMA/CD 在局域网中访问和传输数据，假如一个节点要发出一条消息，那么首先检查通信信道，看是否正在进行数据传输。如果没有，就可以使用线路，发送自己的消息，否则必须等到信道空闲。如果两个节点同时试图发出消息，

就会发生冲突，此时两个节点各自等待随机长度的一段时间，再重发上一条消息。在最早的 802.3 中，规定传输速率为 10Mbps，使用 50 欧姆的电缆，后来这些都有所改变。因为其传输电缆的原因，很多人称之为以太网，其电缆有以下几种，如表 13-6。

表 13-6：通常的以太网电缆类型

名称	电缆	最大段长	每段最多节点	好处
10Base5	粗同轴电缆	500 米	100	可以作为骨干线路
10Base2	粗同轴电缆	200 米	30	价格最低廉
10Base-T	双绞线	100 米	1024	易于维护
10Base-F	光纤	2000 米	1024	可以用于建筑之间

2） IEEE 802.4：令牌环网，在令牌环网中，有一个特殊的数据包，称为“令牌”，沿着网络从一个节点发送到另一个，该令牌包含着特殊的控制信号，使节点能够知道令牌目前是正携带着一条消息，还是处于空闲状态，可以接收一条消息。如果令牌空闲，同时节点需要发出一条消息，以“捕获”该令牌，并将自己的消息附加其中，随后该令牌就不能再携带其它消息，除非将目前的消息“投递”出去。

3） FDDI：FDDI 是高性能的光纤局域网，其传输速率为 100Mbps，传输距离可达 200KM，其间可以连接 1000 台工作站。FDDI 的使用类同于 IEEE 802 规范的 LAN。但由于其高带宽，可以作为连接铜质局域网的骨干网。

3．WebGIS——万维网地理信息系统

3．1 因特网和 GIS

毋庸置疑，因特网的发展为 GIS 发展带来了极大的便利，同时也 GIS 理论及技术研究提供了新的领域。作为信息系统以及一门学科，GIS 可以从因特网的发展受到的影响主要有以下几个方面：

1）GIS 研究者利用新闻组或者电子邮件进行 GIS 技术问题的探讨。

2）网络远程教育，即教授将教案以HTML文档形式放在网上，学生下载使用，并且可以利用电子邮件进行提问，这样就形成了“虚拟大学（Virtual University）”。GIS的教育也可以以该种方式进行，目前有关GIS教育网址有，网址http://giswww.kingston.ac.uk*，提供了三个与GIS有关的专题；网址http://www.utexas.edu介绍了有关GIS和地理教学面临的挑战，网址http://www.census.gov/pub/geo/gis-faq.txt有许多关于GIS常见问题的问答。

3）GIS 软件的下载，GIS 软件公司可以定期将其开发软件的最新版本放在其站点上，以供用户下载试用。

典型地理信息系统方面 WWW 网址介绍[宫鹏]

（1）GIS WORLD INC. *http://www.gisworld.com*

《地理信息系统世界》是目前国际有关 GIS 方面最畅销的杂志，其网址上列出了该公司出版的任何一期的内容目录，用户还可以利用搜索器进行内容搜索。

（2）GIS 软件开发商网址

ESRI 公司(*http://www.esri.com*)，Intergraph 公司(*http://www.intergraph.com*)，MapInfo 公司(*http://www.mapinfo.com*)。

* 本文列出的一些URL极有可能在读者看到时已经“失效”，这几乎是Web上司空见惯的事情了。

这些被认为是一流的GIS网址，有丰富的GIS信息，包括的内容有：公司介绍，产品，公司出版物，年度报告，会议预报，用户信息，产品订购，技术咨询，程序设计技巧，就业机会，产品展示和免费软件下载等等。

（3）虚拟旅行家网 *http://www.vtouris.com*

该网址允许用户进行交互的地图显示，首先是世界地图，然后是各国或各地区地图，最后是城市地图。其地图是栅格图像，利用空间位置实现用户交互操作，有大量的文本、图像等描述信息供用户查阅。

（4）加拿大地球信息科学网 *http://abbott.ccm.emr.ca*

该网站是一个国家级的对地球信息进行管理的机构，它提供大量的产品和服务，包括加拿大的卫星图像，国家地图集数据，加拿大地名数据库，地形数据库，道路网络，数字地形数据等等。此外还提供了获得以上数据的服务，用搜索器得到搜寻地区的数据。

（5）填充式数据查询和表达网址 *http://tiger.census.gov/cgi-bin/mapbrowse-tbl*

该网站采用填充式(fill-in-the-blanks)的方式进行查询。它提供了一个表格，用户填写有关信息后，将请求传入服务器端的GIS、制图、数据库查询等软件运行相应的搜索和查询，也可以是空间分析等，其结果可以是地图或者表格，并传回浏览器。微软的门户站点MSN也提供了类似的功能 *http://msn.maps.expedia.com/quickmaps.asp*。

（6）GIS论坛 *http://gis.itc.nrcs.usda.gov/gishome.html*

该网站包括GIS数据以及软件的发布，GIS领域内的事件，一些技术问题的探讨，GIS书籍的介绍等等，此外可以通过Email的方式进行相关问题的讨论。

4）空间数据发布和下载，数据是GIS系统中最为重要的部分，数据的录入和预处理也是GIS应用开发过程中耗费时间、资金最多的一个环节，而通过因特网实现数据共享，可以降低GIS工程的开发成本。由于因特网的迅速发展，促进了电子商务的兴起，空间数据当然也可以作为一种特殊的商品在因特网上发售。与后面提及的WebGIS方式相比，这里的数据下载还主要是利用文件传输的方式实现。由于因特网上信息量浩大，常常使得找到真正需要的数据成为一件困难的事情，而应用空间元数据可以使用户迅速定位需要的数据并进行下载。1994年美国政府开始发展国家空间数据基础设施（NSDI），通过确定元数据标准，要求各级政府机构采用元数据的方式在网络上对其所生产的数据进行描述，达到各机构间数据生产和共享的目的。

5）此外，由于因特网的发展，打破了传统的时间——空间联系方式，形成了空间事物的新的组织形式，称为计算机网络信息空间(Cyber Space)，是目前人文地理学研究中的热点，也将是GIS探讨的重要课题。

3．2WebGIS简介

Web技术和GIS技术相结合，最为激动人心的产物就是WebGIS(万维网地理信息系统)。WebGIS，简言之，就是利用Web技术来扩展和完善地理信息系统的一项新技术。由于HTTP协议采用基于C/S的请求/应答机制，具有较强的用户交互能力，可以传输并在浏览器上显示多媒体数据，而GIS中的信息主要是需要以图形、图像方式表现的空间数据，用户通过交互操作，对空间数据进行查询分析。这些特点，就使得人们完全可以利用Web来寻找他们所需要的空间数据，并且进行各种操作。具体地讲，WebGIS的应用可以分为以下几个层面：

1）空间数据发布

由于能够以图形方式显示空间数据，较之于单纯的FTP方式，WebGIS使用户更容易找到需要的数据；

2）空间查询检索

利用浏览器提供的交互能力，进行图形及属性数据库的查询检索；

3）空间模型服务

在服务器端提供各种空间模型的实现方法，接收用户通过浏览器输入的模型参数后，将计算结果返回。换言之，利用 Web 不仅可以发布空间数据，也可以发布空间模型服务，形成浏览器/服务器结构(Browser/Server，B/S)。

4）Web 资源的组织

在 Web 上，存在着大量的信息，这些信息多数具有空间分布特征，如分销商数据往往有其所在位置属性，利用地图对这些信息进行组织和管理，并为用户提供基于空间的检索服务，无疑也可以通过 WebGIS 实现。

与传统的地理信息系统相比，WebGIS 有其特殊之处，主要表现在：

1）它必须是基于网络的客户机/服务器系统，而传统的 GIS 大多数为独立的单机系统；

2）它利用因特网来进行客户端和服务器之间的信息交换，这就意味着信息的传递是全球性的；

3）它是一个分布式系统，用户和服务器可以分布在不同地点和不同的计算机平台上。

3．3WebGIS 的实现技术

WebGIS 是网络 GIS 的一个重要组成部分，网络 GIS 的一些概念，如客户机/服务器模式、分布式数据管理等，也可以应用于 WebGIS，但是在 WebGIS 实现时，还要着重考虑两个问题，即控制网络传输数据量以及必须通过浏览器与用户进行交互。

目前已经有多种不同的技术方法被应用于研制实现 WebGIS，包括 CGI(Common Gateway Interface,通用网关接口)方法、服务器应用程序接口(Server API)方法、插件(Plug-ins)法、Java Applet 方法以及 ActiveX 方法等等，下面对这些技术进行简单的描述和比较。[夏福祥]

3．3．1 CGI 方法

CGI 是一个用于 Web 服务器和客户端浏览器之间的特定标准，它允许网页用户通过网页的命令来启动一个存在于网页服务器主机的程序（称为 CGI 程序），并且接收到这个程序的输出结果。CGI 是最早实现动态网页的技术，它使用户可以通过浏览器进行交互操作，并得到相应的操作结果。

利用 CGI 可以生成图像，然后传递到客户端浏览器（目前大多数主页的访问者计数器就是采用 CGI 程序实现的）。这样，从理论上讲，任何一个 GIS 软件都可以通过 CGI 连接到 Web 上去，远程用户通过浏览器发出请求，服务器将请求传递给后端的 GIS 软件，GIS 软件按照要求产生一幅数字图像，传回远程用户。

实际上，由于设计的原因，大多数 GIS 软件不能直接作为 CGI 程序连接到 Web 上，但是，有以下的两种技术比较成功。

1）用 CGI 启动后端的批处理制图软件，这种软件的特点是用户可以直接在计算机终端一行一行地输入指令来制图。其特点是用户的每一个要求都要启动相应的 GIS 软件，如果软件较大，启动时间就会很长。

2）CGI 启动后端视窗(Windows)GIS 软件，CGI 和后端 GIS 软件的信息交换是通过“进程间通讯协议（IPC-Inter Process Communication）”来完成，常用的 IPC 有 RPC(Remote

Procedure Call)和 DDE(Dynamic Data Exchange)。其优点在于，由于 GIS 软件是消息驱动的，CGI 只要通过发送消息，驱动 GIS 软件执行特定操作即可，不需要每次重新启动。

3．3．2 Server API 方法

Server API 类似于 CGI，不同之处在于 CGI 程序是单独可以运行的程序，而 Server API 往往依附于特定的 Web 服务器，如 Microsoft ISAPI 依附于 IIS(Internet Information Server)，只能在 Windows 平台上运行，其可移植性较差。但是 Server API 启动后会一直处于运行状态，其速度较 CGI 快。

3．3．3 插件方法

利用 CGI 或者 Server API，虽然增强了客户端的交互性，但是用户得到的信息依然是静态的。用户不能操作单个地理实体以及快速缩放地图，因为在客户端，整个地图是一个实体，任何 GIS 操作，如放大、缩小、漫游等操作都需要服务器完成并将结果返回。当网络流量较高时，系统反应变慢。解决该问题的一个办法是利用插件技术，浏览器插件是指能够同浏览器交换信息的软件，第三方软件开发商可以开发插件以使浏览器支持其特定格式的数据文件。利用浏览器插件，可以将一部分服务器的功能转移到客户端，此外对于 WebGIS 而言，插件处理和传输的是矢量格式空间数据，其数据量较小，这样就加快了用户操作的反应速度，减少了网络流量和服务器负载。插件的不足之处在于，象传统应用软件一样，它需要先安装，然后才能使用，给使用造成了不方便。

下面是利用 WebGIS 插件在因特网上显示空间数据的例子，利用插件能够方便地对地图进行漫游、缩放和查询操作。

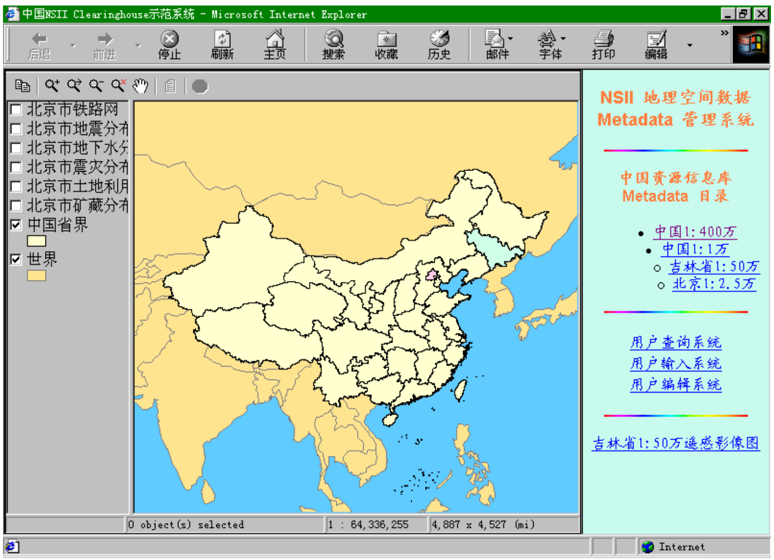


图 13-15：WebGIS 插件，来自于北京大学数字地球工作室
(HTTP://WWW.CYBERGIS.ORG.CN)

3．3．4Java Applet（Java 小应用）方法

WebGIS 插件可以和浏览器一起有效地处理空间数据，但是其明显的不足之处在于计算

集中于客户端，称为“胖客户端”，而对于 CGI 方法以及 Server API 方法，数据处理在服务器端进行，形成“瘦客户端”。利用 Java 语言可以弥补许多传统方法的不足，Java 语言是一种面向对象的语言，它的最大的优点，就是 SUN 公司提出的一个口号“写一次，任何地方都可以运行(Write once, run anywhere.)”，即指其跨平台特性，此外 Java 语言本身支持例外处理、网络、多线程等特性，其可靠性和安全性使其成为因特网上重要的编程语言。

Java 语言经过编译后，生成与平台无关的字节代码(Bytecode)，可以被不同平台的 Java 虚拟机(JVM-Java Virtual Machine)解释执行。Java 程序有两种，一种可以独立运行，另一种称为 Java Applet，只能嵌入 HTML 文件中，被浏览器解释执行。用 Java Applet 实现 WebGIS，优于插件方法的方面是：1）运行时，Applet 从服务器下载，不需要进行软件安装；2）由于 Java 语言本身支持网络功能，可以实现 Applet 与服务器程序的直接连接，从而使数据处理操作既可以在服务器上实现，又可以在客户端实现，以实现两端负载的平衡。图 13-16 是利用 Java Applet 实现的 WebGIS 系统结构。

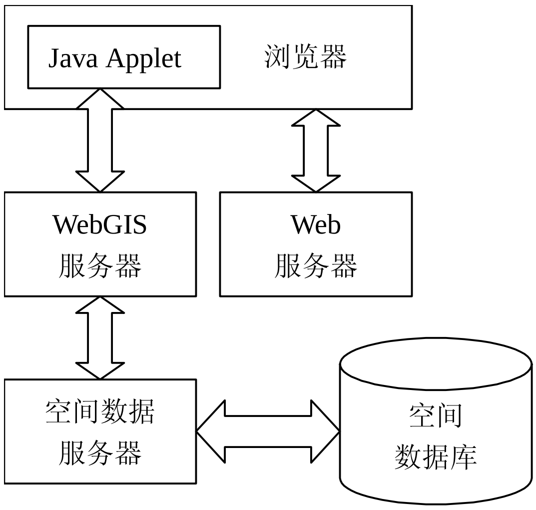


图 13-16：利用 Java Applet 实现的 WebGIS 系统框架

3．3．5ActiveX 方法

另一项可以实现 WebGIS 的技术是 ActiveX，它是在微软公司 OLE 技术基础上发展起来的因特网新技术，其基础是 DCOM(Distributed Component Object Model)，它不是计算机语言，而是一个技术标准。基于这种标准开发出来的构件称为 ActiveX 控件，可以象 Java Applet 一样嵌入到 HTML 文件中，在因特网上运行。与 Java Applet 相比，其缺点是只能运行于 MS-Windows 平台上，并且由于可以进行磁盘操作，其安全性较差，但是优点是执行速度快，此外由于 ActiveX 控件可以用多种语言实现，这样就可以复用原有 GIS 软件的源代码，提高了软件开发效率。

上面描述了几种 WebGIS 的实现方案，在实际的系统建设中，可以根据待发布数据的数据量、数据类型、Web 服务器软件、客户端的要求等确定采用不同的方案，选择相应的软件。

表 13-7：国外几个重要的WebGIS技术特征[宋关福等]

	MapInfo ProServer	**GeoMedia Web Map**	**Internet Map Server(IMS)**	**MapGuide**	**ModelServer/ Discovery**
公司	MapInfo Corp.	Intergraph Corp.	ESRI Inc.	Autodesk Inc.	Bently
服务器操作系统	Windows NT/95	Windows NT	Windows NT	Windows NT	Windows NT

Web 服务器	支持 CGI 的 Web Server	Internet Information Server	Internet Information Server 或 Netscape Server	支持 CGI 的 Web Server	Netscape Server
其他服务器端软件	ODBC 、MapInfo 4.x 、MapBasic	ODBC	ArcView 或 MapObjects 应用、ODBC	ODBC	MicroStation GeoGraphics ODBC
客户端操作系统	Windows 系列 、Macintosh 、UNIX	Windows NT/95	Windows 系列 、Macintosh 、UNIX	Windows NT/95	Windows 系列、Macintosh、UNIX
客户端浏览器	支持 HTML 的任意浏览器	Internet Explorer, Netscape Navigator	支持 HTML 的任意浏览器	Internet Explorer ,Netscape Navigator	Internet Explorer, Netscape Navigator
客户端是否需要插件(plug-in)/控件(control)	不需要	如果使用 Netscape Navigator 浏览器，需要安装 ActiveCGM 插件；如果使用 Internet Explorer 浏览器，会自动下载 ActiveCGM 控件	自动下载 Java Applet 或者 ActiveX 控件	需要安装 MapGuide 插件(1 兆左右)	需要安装 VRML、CGM、SVF 等插件
网络传递的图形格式	JPEG(栅格图)	ActiveCGM(栅格图和矢量图)	JPEG/GIF(栅格图)	MWF(矢量图)	JPEG,PNG, VRML,CGM,SVF(栅格图和矢量图)
地图预出版处理	动态生成地图	动态生成地图	动态生成地图	需地图预出版处理	动态生成地图
可发布的数据格式	MapInfo 地图文件	MGE 工程、MicroStation DGN 文件,FRAME 文件 ,MGEDM 文件,ArcView Shape 文件 ,Arc/Info Coverage,SDO 文件	ArcView Shape 文件 ,Arc/Info Coverage,SDE 地图文件，Autodesk DWG 文件	Autodesk DWG	GeoGraphics 工程文件,MicroStation 设计文件

目前，WWW 领域，可扩展标记语言（Extensible Markup Language，XML）得到了越来越多的重视，它可以成为一种“元语言”，用于定义特定领域的标记语言，同样在空间信息的 Internet 发布中，也可以采用 XML 来定义地理信息的特定语言标记，以容易而一致的方式格式化和传送数据。

第十四章 地理信息系统应用实例

导读：本章介绍了地理信息系统在一些具体领域的应用。

由于 GIS 是用来管理、分析空间数据的信息系统，所以几乎所有的使用空间数据和空间信息的部门都可以应用 GIS。由于各个部门的不同，GIS 在具体业务系统中所占的比重、应用方式也各异，结合比较紧密的如城市规划、环境领域等，相对松散的有商业、医疗卫生领域。

本章简单介绍了一些 GIS 应用实例，针对不同实例，介绍方法也不同，有的概述一个具体区域的应用，有的叙述工作流程，有的介绍了工作原理和分析方法，可以对相关领域的 GIS 建设提供借鉴，也可以作为其它领域建设 GIS 的参考。

本章也引入了一些新的概念，如 AM/FM、GIS/T 以及一些具体的分析方法，如动态分段模型，它们处于 GIS 与具体领域的交叉部分，同时也是 GIS 理论和技术的有益补充。

1．城市规划、建设管理

城市是人类活动高度集中的区域，同时也是信息、物质高度集中的区域。随着科技的进步和经济的发展，城市系统越来越复杂，数据和信息越来越多，服务要求越来越高。城市管理面临着新的挑战，为了城市的现代化、生态平衡和持续发展，城市需要全面的规划，而地理信息系统给城市的规划和管理带来了新的工具。

在城市管理中，可以应用GIS的方面非常多，如土地、道路、管网、环境、人口等等诸多要素都可以通过地理信息系统进行管理*，并且形成了应用于专门领域的GIS工具，如土地信息系统（LIS-Land Information System）和AM/FM（自动制图/设施管理，Automatic Mapping/Facility Management），前者侧重于管理地块的位置、面积、权属以及地籍的动态变化；后者则主要对电力、通讯等设施的管理和分析、制图。

城市建设规划涉及的因素非常多，开发新城要征用土地，改建旧城要拆迁安置，同时需要基础设施、公共服务设施的配套。在开发建设活动中，如果不注意各工程项目之间的协调，就可能造成混乱，而采用 GIS 对各种信息进行管理，并基于此进行分析和辅助决策，可以有效地防止这种混乱局面的出现。由于城市在不断地建设发展，所以需要随时更新城市基础数据库，这就要求应用 GIS 管理日常城市建设活动，以保证信息的时效性。下面介绍 GIS 在密尔沃基（Milwaukee）市城市建设日常管理方面的应用。

美国威斯康星州的密尔沃基从 1976 年开始建设城市地理信息系统，并不断发展、完善。由于数据库比较完整，地理信息系统在很多方面发挥了作用，主要可以分为三个方面的活动。

1）地图更新

密尔沃基市常用地图分为三种：

第一种是地籍房产图（当地称为 Plot Map），由政府税务部门负责编制。这种地图通常在以下情况下需要修改更新：

- 地契内容改变，包括地块边界的变化以及地块权属的变化；
- 在一个大的地块内再划分为若干小块，产生新的土地所有者；
- 公共道路的拓宽或者连通；

---

* 下面提及的GIS在环境、道路等方面的应用也可以认为是地理信息系统在城市管理中应用的部分。

- 小地块合并成为大地块；
- 其它变化，如道路名称、地名的改变以及错误的纠正。

第二种是一般的测绘图（当地称为 Quarter-Section Map）。这种图由政府市政工程部门负责编制，图面上标有主要的地块边界（不作为法律依据），但是没有普通的建筑物，通常在下列情况下需要更新：

- 城市建设造成地面重要物体的边界变化；
- 重新测量，纠正过时的内容；
- 道路打通、拓宽或封死；
- 大的地块合并或重新划分；
- 街道名称改变；
- 地图上的错误纠正。

第三种是土地使用图（当地称为 Land Use Map）。这种地图反映土地的实际用途，也反映了主要的地块边界。土地使用图由城市规划部门负责编制，主要为城市规划服务，通常在以下条件下需要更新：

- 房屋的拆除和重建；
- 房屋改变用途；
- 地块边界改变；
- 地块重新划分；
- 地块兼并；
- 道路变化；
- 道路名称的修改；
- 地图错误纠正。

以上三种地图由三个相互独立的部门各自负责，在信息上存在很多重复内容，在地图更新时也需要大量的重复工作。在建立城市地理信息系统之后，三种地图上各类信息的更新工作明确地分配到三个部门，三种数字化地图集中地存放于数据库中，实现数据共享，每个部门对数据的修改内容可以被其它部门得到，减少了数据冗余和重复工作。

2）土地区划管理

在美国，大多数地方政府都用土地区划（Zoning）来指导、限制城市土地利用和城市发展，但是具体做法各地有许多差异，在密尔沃基市，先由城市规划部门编制土地使用规划，经批准后再制定区划，区划的形式为文本和图件，它是城市建设和审批修建申请的依据。

密尔沃基市的土地区划图分为用途图、范围图和高度图，用途图用于限制和规定相冲突的土地利用，如不准贴近主要交通干道修建大型商业设施等；范围图对建筑之间的距离以及建筑后退道路等作出了规定；高度图则规定了建筑本身的高度、相邻建筑在高度上的相互关系的限制。

在采用 GIS 后，这三种区划图分为三个图层进行管理，并可以叠加在一起显示，避免各个部门数据不一致的情况，此外，可以很方便地检查规划图的错误，避免法律上的纠纷；也便于税收部门查询地产信息，调整对土地所有者的税收。

3）建筑审批处的内部工作管理

每个建筑物在修建之前，必须经过政府主管部门批准，由于申请非常多，并且需要实地勘察，工作繁重，如何分配并确定每个经办人员的工作，以充分调动每个人的积极性成为一个难题。在建立地理信息系统后，修建项目申请卡片被记录到属性数据库中，其中的建筑物地址属性与地块空间数据相关联，这样可以很方便地将一定时期内所有的修建申请的位置分布显示在地图上（图 14-1），可以方便的根据申请量的多少、到市中心的距离、前往勘察的边界程度，划分每个工作人员的负责范围，提高了工作效率，改善了建筑审批处的内部管理。

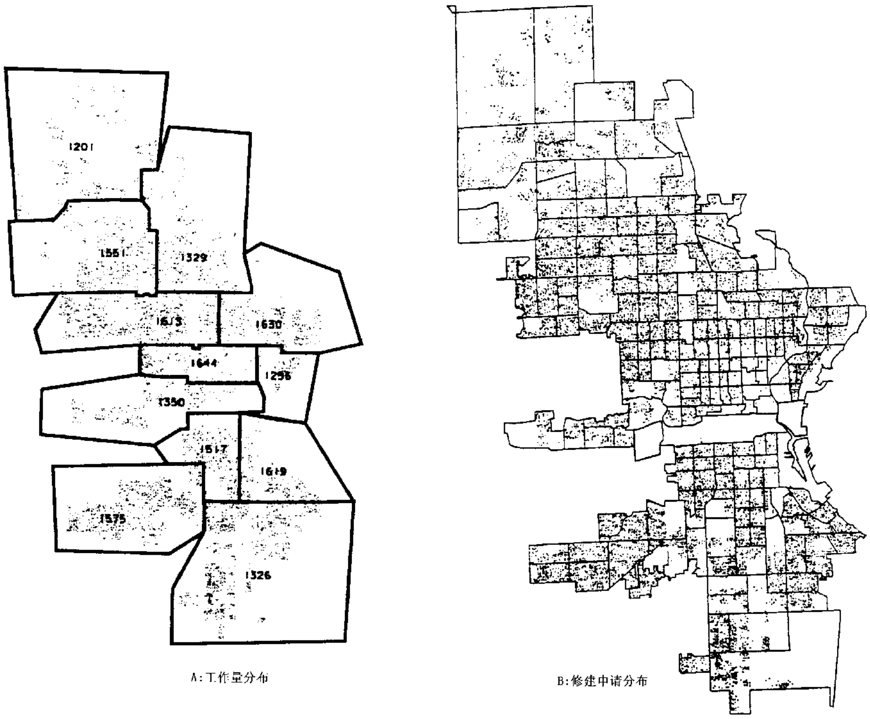


图 14-1：利用 GIS 进行建筑审批工作协调

2．农业气候区划

2．1 项目背景与需求

自八十年代以来，中国农业生产环境、气候环境发生了巨大变化，包括：

1）农业科学技术进步，高新技术引进，农业朝高产、优质、高效、低耗方向发展；种植业向粮食—经济—饲料作物三元结构转变；小生产、大市场向规模经营方向发展，逐步实现种养加、产供销、贸工农一体化。农村产业结构调整向农业气候区划提出了新的要求。

2）气候条件与气候资源本身发生了变化。如东北地区平均气温升高，北方地区干旱范围扩大，长江流域洪涝增多，特别是九十年代以来，异常气候事件呈现明显增多的趋势。有必要重新认识气候资源的变化及其合理利用和保护问题。

3）农业气候区划技术条件发生了巨大变化。“3S”（GIS、RS、GPS）与网络技术在农业气候资源动态监测与开发应用中展示出广阔前景。

4）农业生产的发展，对气候资源开发利用和保护提出了更高要求。气象部门必需依靠科技进步，技术创新适应社会的需求。建立一套基于“3S”、网络平台的区划信息系统，有助于广大气象台站在气象服务手段和技术上得到提升。

5）面对 21 世纪农业新技术革命和可持续农业战略，农业气候资源作为一种重要的投资环境，有必要进行科学、客观的评价。

中国气象局提出建设“第三次农业气候区划”项目的目的是：采用新技术、新方法、新资料，开发“农业气候区划信息系统（Agriculture & Climate Distributed Information System, 简称 ACDIS）”软件，建立气候资源开发利用和保护监测体系，实行资源平面与立体，时间与空间全方位优化配置；发挥区域气候优势，趋利避害减轻气候灾害损失，提高资源开发的

总体效益。为各级政府分类指导农业生产，农村产业结构调整，退耕还林防止水土流失等提供决策依据，为地方政府服务。

2．2 农业气候区划信息系统（ACDIS）系统结构及工作流程

2．2．1 系统结构

农业气候区划信息系统（**ACDIS**）是基于 GIS 工具平台，建立起面对专业技术人员的专用工具，适用于农业气候资源监测评价、气候资源管理与分析，小网格气候资源推算与空间查询、省地县三级区划产品制作等。具体实现了以下功能：

1）地理基础信息管理：管理工作区的基础地理数据，如行政区划，水系，交通数据等；

2）小网格资源信息管理：管理栅格格式的、与农业气候区划有关的信息，包括各种 DEM 数据；

3）小网格资源推算与区划产品制作；

4）农业气候资源监测与评价；

5）农业气候区划成果演示。

2．2．2 系统工作流程

根据“3S”等新技术在区划中应用需求，建立了农业气候区划工作基本流程（如图 14-2）：

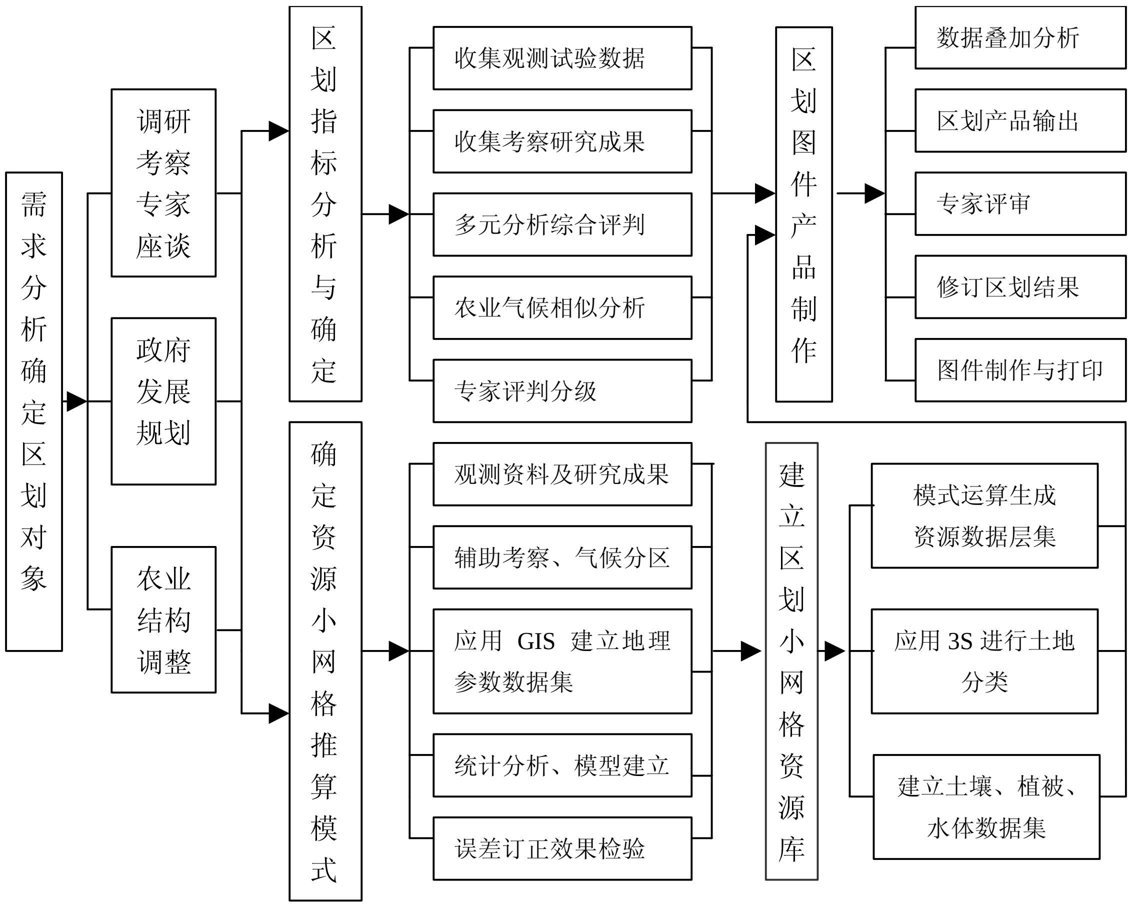


图 14-2：农业气候区划工作基本流程图

2．3“3S”在项目中的应用

1）利用 GIS 对农业气候区划综合要素空间查询和管理

ACDIS 利用 GIS 工具提供的基本功能，对其它子系统输入、处理和生成的气候资源等数据进行综合查询和管理。生成不同查询条件下的区划产品，作为区划专题内容，进而实现对区划各类产品的矢量图、栅格图、DEM、注记和属性数据进行以地理表达式为条件的逻辑查询以及不同图件和属性数据的综合查询、管理，并把查询结果制图输出到绘图仪或打印机，或保存为其它格式文件。

2）利用 GIS 对气候资源进行小网格推算模式研究

气候资源小网格推算模式研究是区划的一项基础性工作。在收集山区气候研究成果基础上，辅助以气象哨、水文站雨量资料后，通过 GIS 平台可快速计算和获取测点地理参数（高程、坡向、坡度）。采用统计分析方法，根据要素的统计特征值和地理特征将全省划分若干气候区，分别建立各区域气候要素推算统计模式，通过验证和残差订正，应用到气候资源小网格推算中。

3）GIS 在农业气候资源分析中的应用

应用 GIS 来定量采集、管理、分析具有空间特性的气候资源，建立 GIS 分析气候资源的思想、方法、步骤，包括数字高程模型建立、GIS 农业气候资源数据建立、空间分析模型建立、气候资源分析计算、气候分区及定量分析等。

4）利用“3S”提取农业背景信息参与区划计算

农业气候区划是根据农作物生长发育过程中对气候条件的要求和气候资源的地理分布特征来进行分区划片的，在某种农作物的气候可种植区内还有不同的地物类型，不同的农作物要求不同的地理环境。为使农业气候区划对农业生产更具有指导作用，将非气象因子引入到农业气候区划中。农业气候区划对象中往往对土壤 PH 值要求很高，根据土壤类型分布可以得出土壤 PH 值的分布，将其作为区划的一个关键指标，使得区划更加有实际应用意义。利用 GIS 将土壤分类图作为一项数据层参与气候资源数据层集运算，得出包含土壤类型信息的区划结果。

江西省在全省优质早稻种植气候区划和万安县脐橙种植综合区划中（图 14-3），除了应用 1∶25 万的地理数据，还结合了 TM 影像数据，辅助 GPS 定位抽样，把早稻、脐橙的可能种植区（农田、荒山荒坡）提取出来，排除了山体、水体、居民点、道路等不能种植脐橙和早稻的区域，把可能种植区与农业气候区划图做逻辑交集运算，得到了全省优质早稻和万安县脐橙种植规划图。

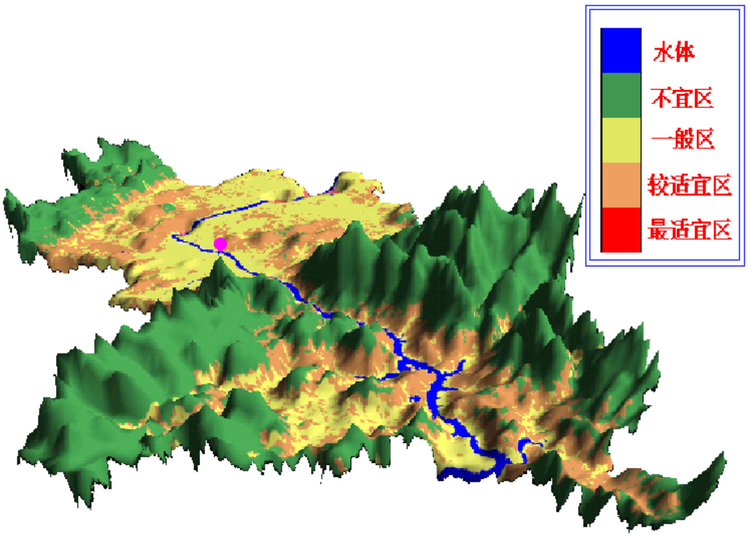


图 14-3：江西省万安县脐橙种植区划

5）利用 GIS 建立集区划—资源动态监测—高产栽培技术为一体的信息服务系统

利用 GIS 工具，开发出基于 Web 的、对气候条件敏感的两系杂交稻制种气候资源空间配置信息系统，将区划服务于作物生产基地选择、关键生育期气象服务、农业生产技术指导等各个方面。

3．大气污染监测管理

随着经济的发展，环境污染直接影响了人们的生活质量，环境质量问题也得到了越来越多的重视。污染环境包括水污染、大气污染、固体废弃物污染等，其中就大气污染而言，城市区域由于受到工业生产、居民生活的影响，成为大气污染发生的集中区域，历史上几次严重的污染事故，如伦敦烟雾事件（1952）、洛杉矶光化学烟雾事件（1943），都是发生在大城市。近几十年来，研究者对大气污染问题进行了大量研究，并且通过实验或计算来建立适合于特定区域的大气污染物扩散模式以及确定相关参数的计算方法。

城市大气污染的来源主要包括点状污染源和线状污染源，前者主要包括烟囱，后者则指汽车尾气排放。而污染的扩散的影响因子，除了排放量、排放物之外，还受到气象条件、下垫面等因素的影响。通过大气动力学的研究，研究者给出了用来描述污染物在大气中扩散规律的各种解析方程。这些大气扩散公式为空气污染预报提供了一定的基础，可以为现有污染源条件在不利的气象条件下减少有害污染物的排放、在城市规划中合理进行区域规划以及环境影响评价提供了切实有效的信息。

无论是点源污染，还是线源污染，其空间分布以及属性可以通过地理信息系统进行管理，而污染扩散的影响因子的空间分布同样可以作为 GIS 的空间数据组成部分，所以，基于 GIS 可以建立大气污染扩散模型，进而，GIS 也提供了丰富的功能以表现污染物强度空间分布，可以查询强度分布状况，并可以结合其它社会经济数据，进行更加细致的评价分析。

下面介绍内蒙古自治区包头市在利用 GIS 进行大气污染扩散模拟中的应用。

包头市是中国最大的稀土工业基地和著名的钢铁机械工业基地之一，是门类齐全，体系较为完善的现代化工业城市。包头市工业能源消耗以燃煤为主，占 56%，其次为焦碳和电力，在能源结构中，重中污染能源占 70.8%，包头工业排放的 SO_2、烟粉尘和氟化物都较大，造成包头市煤烟型和氟污染的特点。其中 21 家重点空气污染源占全市工业排放的 95%，85%，96%（分别为 SO_2、烟粉尘和氟化物）。其中包钢是最大污染厂家，其 SO_2、粉尘、氟化物的排放量高居榜首。

从自然条件上看，包头市地处内陆，属内陆半干旱、中温带大陆性季风气候，全年干燥，无霜期短。年均气温 6.5℃，取暖期 6 个月。由于受地形影响，冬季包头多北风和西北风，夏季多东南风，各月夜间均盛行 NWW 风。近年来城市建筑的快速发展及自然植被，人工绿化面积的不断扩大，市区平均风速呈逐年下降趋势。1995 年年均风速为 2.2m/s。一年中，春季风速最大，秋季风速最小，冬季风速居中。除七月份主导风向与最大风速风向恰好相反外，常年盛行风向与最大风速风向重合。

逆温层影响着污染物的扩散、稀释，包头市冬季逆温频率很高，夏季逆温频率较低，逆温的强度与厚度，冬季均明显大于夏季。

包头是我国大气污染治理的重点城市，包头关于大气污染扩散的研究工作较多。冶金部建筑研究总院 1982 针对包钢地区的烟气综合治理规划，利用风洞模拟试验、现场实验等提出了“大气输送气候学模式”（ATCM）。1989 年包头市环境监测站，针对包头新市区大气扩散模式和 SO_2 容量计算，提出了基于美国 EPA 的 ISC（工业复合源大气扩散模式）的城市多源高斯模式。这些模式的建立为包头市的大气污染治理和管理提供了可靠的依据。

基于 GIS 建立大气污染扩散模型，可以

1）模拟污染物的空间分布，评价不同区域的环境质量；

2）将污染物空间分布与人口密度空间进行复合分析，确定受污染影响的人口数目；

3）预测在给定气象条件下污染物的空间分布；

4）确定不同点源对整个研究区污染总量的贡献；进而

5）为污染整治，如降低排放量、甚至关闭某些污染源，提供决策依据；

6）如果要增加污染点源，可以比较不同的方案（如烟囱的位置，高度等），从中选择最优方案；

7）在城市规划时，作为确定不同用地（居住、工业、商业等）的分布。

在包头市的研究工作中，利用 1 月份平均风速、风向、频率，并将其换算为风频表，对包头市的 37 个高架点源造成的地面 SO_2 浓度的空间变化进行模拟，结果如图 14-4 所示。

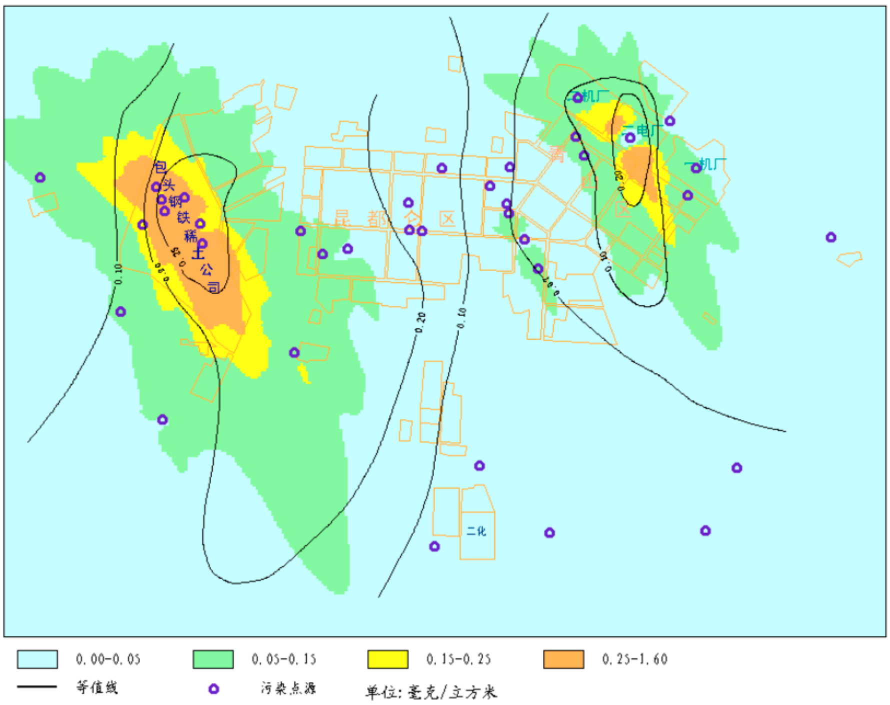


图 14-4：包头市大气二氧化硫分布图

将模型预测结果图与包头市环境监测站绘制的等值线图相比较。模型在工业区的预测值比较切合实际，在昆都仑区，预测值偏小，原因应该是由于包头地区在一月份特定风向条件

下，工业污染对该区域的影响比较小，相比之下居民取暖燃煤造成的二氧化硫污染就较突出。总体来看，无论从工业区还是居民区，模型预测的二氧化硫的浓度不存在数量级上的差别。

在大气环境规划中，对规划中欲添加的新污染源需要预测新增污染源造成的污染、对新源的进行合理选址、确定新源的排污量大小以及点源烟囱的高度等。在研究中，拟在青年农场（图上三角点位置）增加一个新的污染源，有三种方案，分别为源强为 100000mg/s、烟囱高 150 米；源强 1000000mg/s、烟囱高 150 米；源强 1000000mg/s、烟囱高 200 米。在一月风向风频条件下进行预测，将预测结果叠加在原来的污染分布状况上，结果如图 14-5。

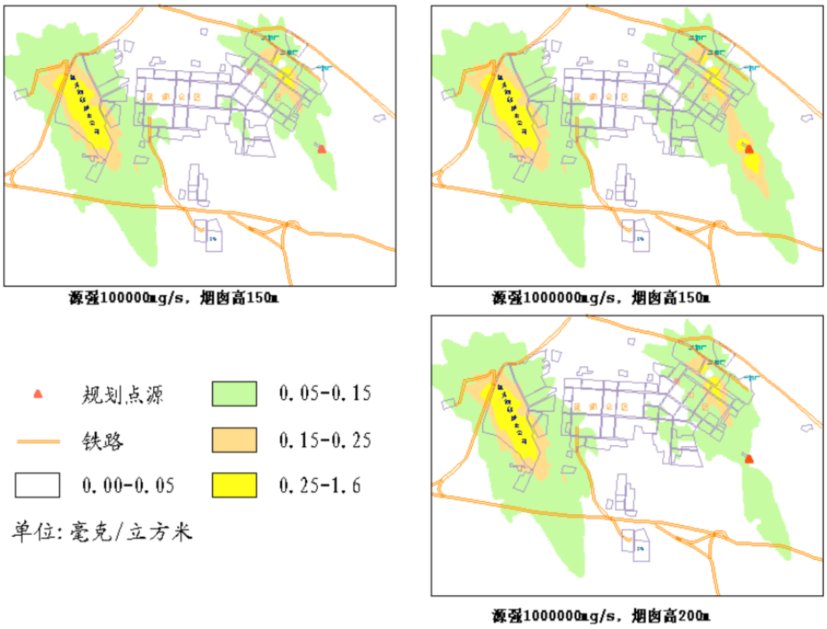


图 14-5：增加不同参数污染源的污染分布图

大气污染的受害人群是大气污染管理和控制中最重要的关注对象，结合包头市新市区（青山区和昆都仑区）的人口分布，研究中分析了添加新源前后大气中 SO_2 污染受害人群的变化。研究方法是将模型的预测结果与包含人口数据的行政区划图进行叠加复合分析，获得各个污染区的面积和人口，结果见表 14-1。

表 14-1：新添加污染源造成的污染人群变化结果表（单位：万人）

浓度单位：mg/m³	0.00-0.05	0.05-0.15	0.15-0.25	0.25-1.6
新源强：100000mg/s 新源高：150 米	46.05	15.33	2.10	0.80
新源强：1000000mg/s 新源强：150 米	36.20	24.30	2.73	1.04
新源强：1000000mg/s 新源强：200 米	41.02	20.02	2.33	0.89
未添加前大气状况	49.83	15.61	2.03	0.80

从图和表可以看出，源强为 1000000mg/s 的污染源造成的高浓度受害人群数量明显增加，而源强为 100000mg/s 的新添污染源造成的高浓度污染受害人群数量变化不大。同时可以看出，增高污染源的高度来降低污染的危害是以低浓度受害人群数量的增加为代价的。从新增污染点源的污染分析来看可以考虑在青年农场位置新建立一个源强为 100000mg/s、烟

囱高为 150 米的新点源。

4．道路交通管理

近年来，GIS 在交通方面的应用得到了广泛的重视，并形成了专门的交通地理信息系统 GIS-T，以满足道路交通管理方面的要求。下面分几个方面介绍 GIS 在公路方面的具体应用。

4．1 路廓设计

路廓设计是公路设计中的一个重要环节，是定出公路最终线向的一个步骤。在路廓设计中，要综合分析多种空间数据，包括大比例尺的土地利用图、地形图以及现有的道路网等。

从各方面收集来的资料合并到 GIS 资料库中，建造约束多边形，确定公路需要避开的区域，通常这些多边形可以分为三级：

1）第一级约束是必须要避开或减至最少；

2）第二级约束是尽量避开或减至最少；

3）第三级约束是可以避开或减至最少。

根据需要可以用权重表示不同的约束等级，并在地图上分别显示，以便设计员在设计路线时尽量绕开这些区域。然后将每条试验性线路输入到 GIS 中，利用缓冲区分析得到路廓多边形，与约束多边形进行叠加分析，计算路廓与每个多边形相交的总面积，并乘以权重，得到该试验线路的约束影响的加权总面积（Total Impact Weighted Area, TIWA），用于比较这些线路设计。

在计算得到 TIWA 之后，可以根据 DEM 进行线向的详细设计，其主要操作是根据 DEM 得到道路横断面，进而进行填挖方的优化，得到最终的路廓设计结果。

4．2 道路管理

4．2．1 道路管理涉及的属性

在公路管理中，通常需要根据其类别、车道数目和路面种类进行分类，其管理的属性包括：

1）公路属性：路网标识，道路标识，路段标识，参考点等；

2）几何属性：路面宽度，路向，车道数目，坡度等；

3）设施属性：中央隔离带和隔离物，护轨，桥梁，交通标志，路灯，交通量计算仪；

4）公用设施：收费所，交叉口，涵洞，交通管理站等；

5）建筑材料属性：面层，覆盖层，基层，路基；

6）道路使用统计属性：交通量，每天平均流量，事故，车重统计，车类统计，进路要求；

7）验收记录属性：强度，粗糙度，障碍物，路面损坏程度，限速区段，车速调查；

8）合同属性：预算，费用，图则，日期等；

9）维修工程属性：工程类别，预算，费用，数量，日期。

这些属性是通过线性定位方法与路网建立关联，通常线性定位的方法有里程碑法和控制

段　法，前者采用路名和公路上的里程点来确定物体的位置；后者则将公路分成相连的、长度不等的控制段，每段有一个编码，并且其属性一致（表 2）。

表 14-2：采用控制段方法记录属性

控制段	表层	车道数目	状态	起点	终点
CS1	砾石	2	坏		
CS2	沥青	4	良好	30.08	30.16
CS3	沥青	2	尚好	30.16	30.2
CS4	沥青	2	良好		

4．2．2 动态分段

交通模型的一个特点是，多种线性物体，重叠在同一个路网上，如限速区段，公共交通线路都与路网重叠，这使得 GIS 在应用于交通时遇到困难，而动态分段（Dynamic Segmentation）方法较好地解决了该问题。

动态分段是在数据库中纪录道路的每种属性的起止点到道路原点的距离，并不是真的将道路切断存储，适合于动态的分析，顾名动态分段。

采用动态分段之后，一个路段（Segment），是路网上两个交点间连线或者弧的一部分，路段的长度用其占连线的比例来表示，具有唯一的标识码。在 GIS 数据库中，路段是依附于路网数据，本身没有坐标。由于采用动态分段将道路的各种属性以及其分布集中在一个图层中进行管理，采用线性定位方法，因而容易实现各种“点线”以及“线线”的叠加查询分析，形如：

交通流量>800,000 AND 路面状况=“良好” AND 道路分级=“省道”

4．3 流量和路径分析

4．3．1 道路网络拓扑

一个路网的表现方式可以偏重其“几何性”或者“拓扑性”，在道路设计中，需要用到其几何表现，而网络分析则着重于拓扑关系（图 14-6）。

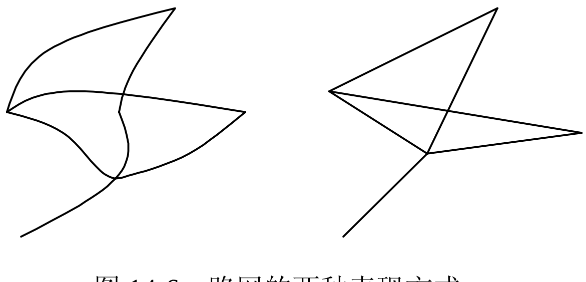
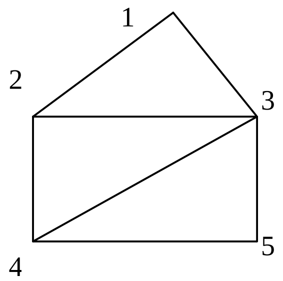


图 14-6：路网的两种表现方式

在交通网络中，大多数线与线之间的交点是具有拓扑性的交点，即在道路的交点处车辆可以转向；但是，由于在 GIS 中，使用的是平面图形（Planar Graph），而有些道路的关系需要在三维空间中才能够正确表现，如立交桥，实际上道路并不直接相交，但是在平面图形上却表现为一个交点，这需要在应用时进行判定，区分拓扑交点和非拓扑交点，以防止出现错误的网络分析结果。

4．3．2 用于道路网络分析的数据结构

进行道路网络分析时，可以只关注其拓扑数据表现，其数据结构和相关算法可以采用已经成熟的“图”数据结构，一种最为简单的实现方式就是“连通矩阵”，记录了结点与结点之间的连接关系（图 14-7）。

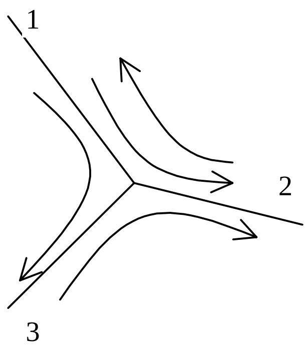


	1	2	3	4	5
1	-	1	1	0	0
2	1	-	1	1	0
3	1	1	-	1	1
4	0	1	1	-	1
5	0	0	1	1	-

图 14-7：网络连通矩阵

在连通矩阵中，单元格数值为 1 表示两个结点直接相连，反之表示不直接相连，这样的连通矩阵可以用于进行连通性、可达性分析；矩阵中每个单元格的数值还可以是结点之间的距离，不直接相连的结点间距离可以用无穷大表示，利用该矩阵可以进行最短路径搜寻。在不考虑连线的方向性，即图是无向的，连通矩阵是对称阵，而在城市道路交通中，常常会出现“单行线”的情况，即连线是有方向的，这样矩阵是非对称的。

在城市交通应用中，还会出现“禁止左转”、甚至“禁止右转”的情况，这时，除了连通矩阵之外，还需要记录每个结点——即路口——的连通属性，同样可以用连通矩阵实现（图 14-8）。



终止道路

起始道路	1	2	3
1	-	1	1
2	1	-	0
3	0	1	-

图 14-8：路口的转向和道路连通矩阵，其中第 2，3 条道路禁止左转

在实际的网络分析中，需要综合利用结点的连通矩阵和结点上道路的连通矩阵。

4．3．3 流量和网络分析的具体应用和相关模型

流量和网络分析一般较多地应用于城市交通中，为城市交通管理和道路规划提供科学的决策支持，其中，有如下四类主要的模型：

1）“出行产生（Trip Generation）”模型

根据交通分析区（Traffic Analysis Zone,TAZ）的土地利用形态以及其它社会经济数据来估计各个区域所产生和吸引的交通量，例如，居住区会产生出行，而商业区则吸引出行。

2）“出行分布（Trip Distribution）”模型

根据出行产生分析的结果，确定各个区域之间的交通量。

3）“交通工具选择（Mode Choice）”模型

将交通量计算分配到交通工具上，计算车流量。

4）“出行分派（Trip Assignment）”模型

根据交通线路的容量和速度，将车流量分配到各条道路上。

上述的四种模型，包含多种传统 GIS 工具所不支持的统计分析运算以及具体的经济地理模型，在具体实现时，可以在 GIS 中运算，将结果输出到其它系统中，进行专业模型计算。

5．地震灾害和损失估计

对地震灾害以及地震次生灾害的评估对于一个区域的降低危险，资源分配以及紧急响应规划具有重要的意义，而通过存储和分析地质构造信息，利用 GIS 可以预测地震发生的“场景”并估计该区域由于地震引发的潜在损失。此外，GIS 也提供了有力的工具使得在地震实际发生时，分析灾害严重程度的空间分布，帮助政府分配紧急响应资源。

进行地震灾害评估时要综合考虑地质构造等各种信息的空间分布，通常包括以下几个步骤（图 14-9）：

1）估计地表震动灾害

需要识别地震源点，然后建立在该点发生地震以及地震波传播的模型，最后根据地表的土壤条件得到最终的震动强度。

2）估计次生的地震灾害

次生的地震灾害包括液化（Liquefaction），滑坡，断裂等，评估这些灾害需要收集相应区域的地质构造信息，计算地表运动的强度和持续时间以及在以前的地震发生过程中这些灾害发生的情况。

3：估计对于建筑物的损害

需要收集地震区域内建筑和生命线的分布状况，然后对每种建筑的建立损害模型，该模型是一个函数，与地表震动强度以及潜在的次要灾害有关。

4：估计可以用金钱衡量和不可以用金钱衡量的损失

可以用金钱衡量的损失包括受损建筑的修复和重建，而不可以用金钱衡量的损失包括人员伤亡。估算这些损失需要相应的社会经济信息，此外，清除垃圾和重新安置费用、失业、精神影响以及其它的长期或短期的影响，需要建立不同的模型，分别加以确定。

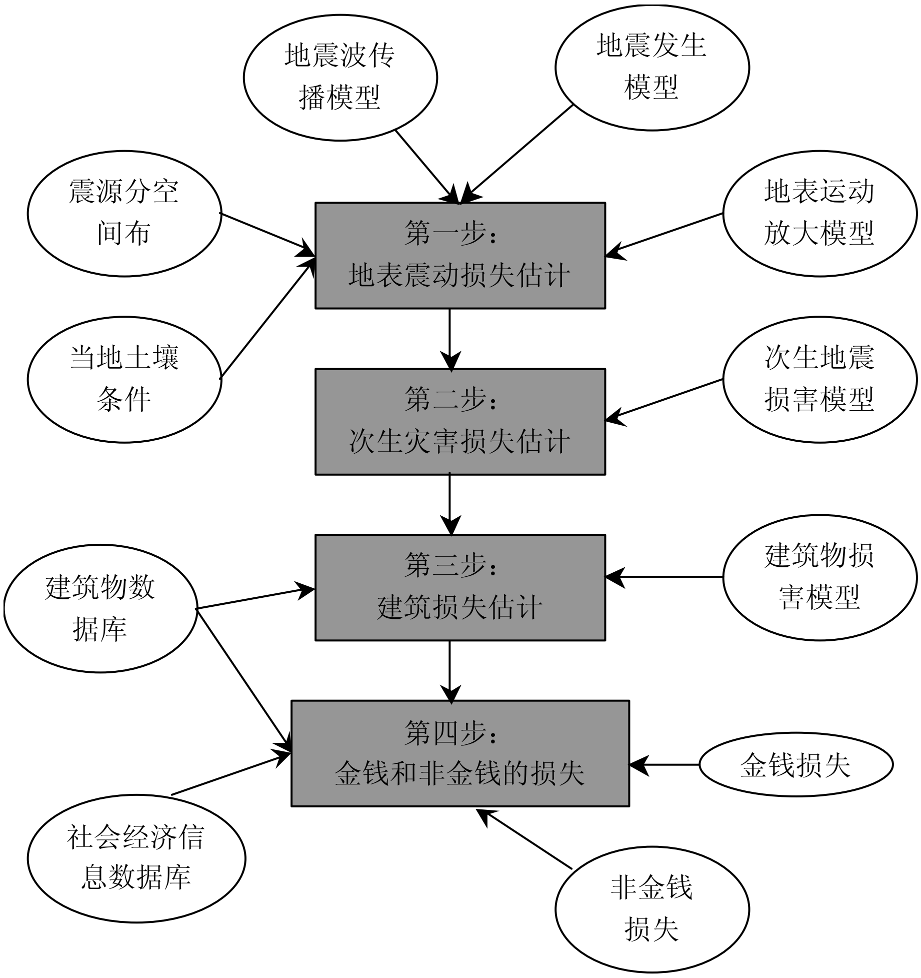


图 14-9：地震损失评估过程

如上图所示，在地震损失评估中，用到了多种空间信息，如地质构造，建筑等等，因此 GIS 是非常理想的进行地震损失评估的工具，图 14-10，13-11 描述了采用 GIS 进行评估的过程。

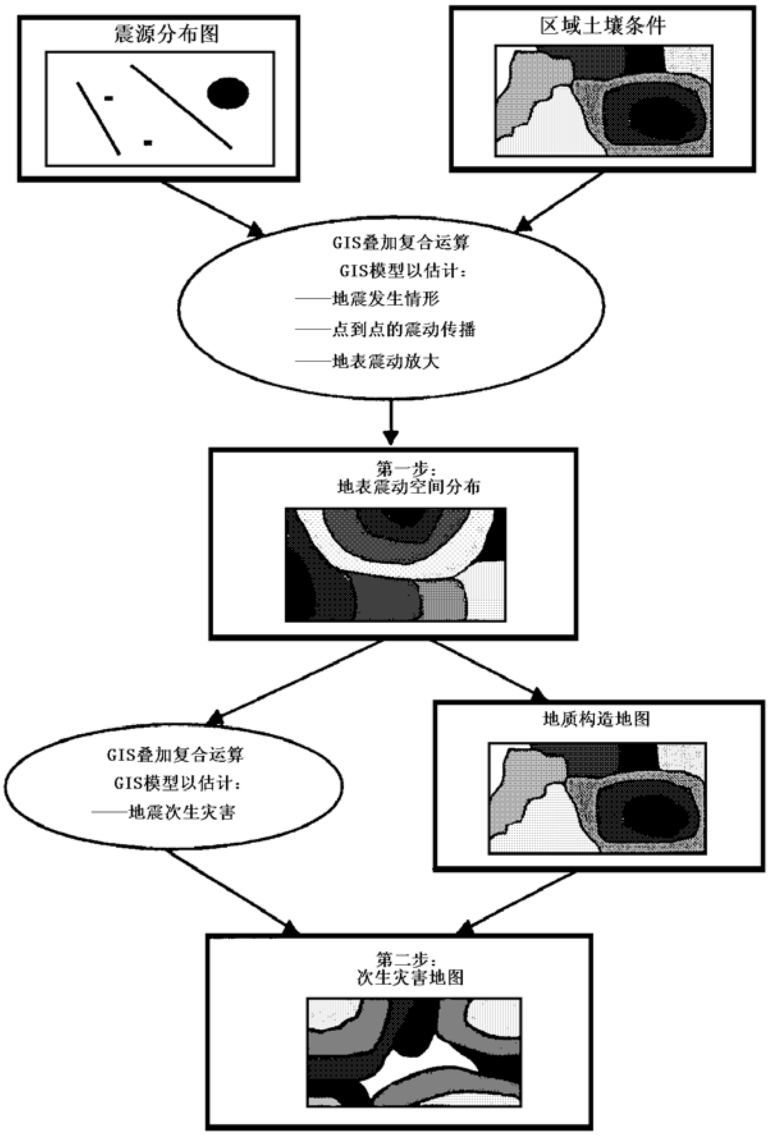


图 14-10：基于 GIS 的地震损失评估过程（1）

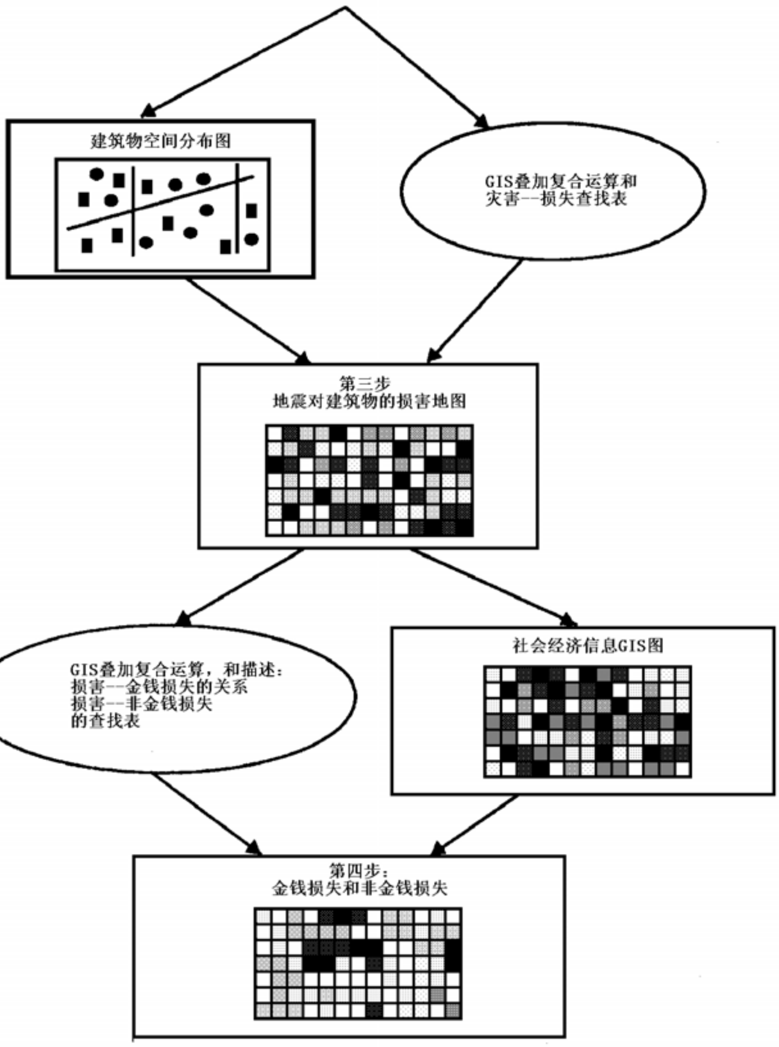


图 14-11：基于 GIS 的地震损失评估过程（2）

通常，地表震动强度可以根据震源位置以及地震波传播公式计算；而次生灾害以及建筑物的损害要根据相关的图件进行计算，并基于上述计算的结果来评估金钱损失和非金钱损失。在分析过程中，由于地震强度以及破坏程度随着到震源的距离增大而衰减，所以要采用缓冲区计算模型；而在计算金钱损失以及非金钱损失时，因为要综合考虑多个因素，要使用叠加复合模型。

6．地貌

地貌学理论发展和生产实践需要加强计量地貌研究。然而，由于地貌现象的复杂性、地貌数据的庞大等多方面的原因，需要在地貌研究中采用 GIS 工具，使其成为地貌定量研究的一个有效途径。

在地貌研究中采用 GIS 技术具体实现步骤和内容归纳为：

1）研究地貌信息内涵和地貌系统的特点。

2）根据地貌信息及其分析方法的特点，结合 GIS 工具建立地貌信息专题分析系统。

3）在地貌信息系统的支持下，建立综合的定量地貌分析模型，利用尽可能多的信息源，分析地貌信息流及其相关物质流、能量流规律，研究地貌形态、物质组成、成因机制、分布特征、发生发展规律及其对人类生产、生活环境的影响作用。

4）开发机制地貌制图系统作为地貌信息系统的输出系统，将以上的分析结果编制成各种平面或立体的地貌分析图件，实现地貌制图自动化，将在资源开发、环境整治、生产建设的实际目标中发挥更大的作用。

6．1 地貌信息的内涵、地貌系统的特点与分析方法

地貌是地壳表层内外营力共同作用到某一发展阶段所形成的三维地表形态，是内、外营力，介质性质和时间的函数，可表述如下：

$$M = F(x, y, z, F_i, F_e, m, t)$$

其中M表示地貌形态；F表示内外营力对地表的作用；x，y，z为空间坐标；F_i表示内营力；F_e表示外营力；m表示构成地貌的介质性质，包括岩性与构造两个方面；t表示作用时间。

内外营力的类型、强度、作用方式、介质性质与作用时间的不同组合，直接影响着地表形态特征。也就是说，在地貌信息中蕴含着归属于不同内外营力和不同发展阶段的特征地貌信息，以此为信息源，提取所需各种特征参数进行复合分析，可以得到不同专题的综合评价结果。

地貌系统是一个开放性的动态系统，是由各种地貌形态类型和要素构成的复杂综合体，它不仅在内部各子系统之间进行着物质、能量和信息流的迁移和转化，而且在系统与外部环境之间也进行着物质、能量和信息流的交换，从而形成一个复杂的物质、能量和信息的传递网络。地貌系统总处于输入——转换——输出的动态过程中，即经常有能量自源区经系统流向耗散区，使系统出现导致有序化的熵减过程，从而具有一定的结构性和相应的熵位，存贮容量和自调节能力。地貌的发展、发生和演化，实质上就是在地貌系统外部环境能量流内、外营力作用下，构成地貌要素的地表物质经历的变形变位（抬升、断错等）分离组合（侵蚀、搬运、堆积等）的复杂物质流传输过程。由于地貌物质传输过程与热力学传导有相似性，许多地貌演化模型，如斜坡演化模型都是热传导方程的形式：

$$\frac{\partial H}{\partial t} = \alpha \frac{\partial^2 H}{\partial x^2} + \beta \frac{\partial^2 H}{\partial y^2}$$

其中 $H=H(x,y,t)$表示地表高程。

由于实际区域地貌系统的过于复杂与庞大，计量研究中往往不能取得数学意义上的精确解释。如以上的演化模型往往是针对某一外力过程，限于极为简单、理想的初始条件并且不考虑介质空间差异的情形；而实际区域内外营力多样化，介质不均一，甚至连初始高程H_0都无法表达成确定的函数式。因此只能采用适当的方式离散化，化无限为有限，求得尽量逼近实况的近似解。

具体地说，就是引入有限元素法与有限差分法的思想，将整个区域地貌体分解成许多“微小”的单元（姑且称之为地貌体，其顶部地表面称之为地貌面），如图 14-12 所示。在小的区域空间内以平代曲，由单个到总体，最终建立整个区域上的种种普适性更强的分析模型以求算各种空间地貌参数和研究复杂地貌过程等问题，可以说离散化是区域地貌信息系统定量分析的前提与特色。

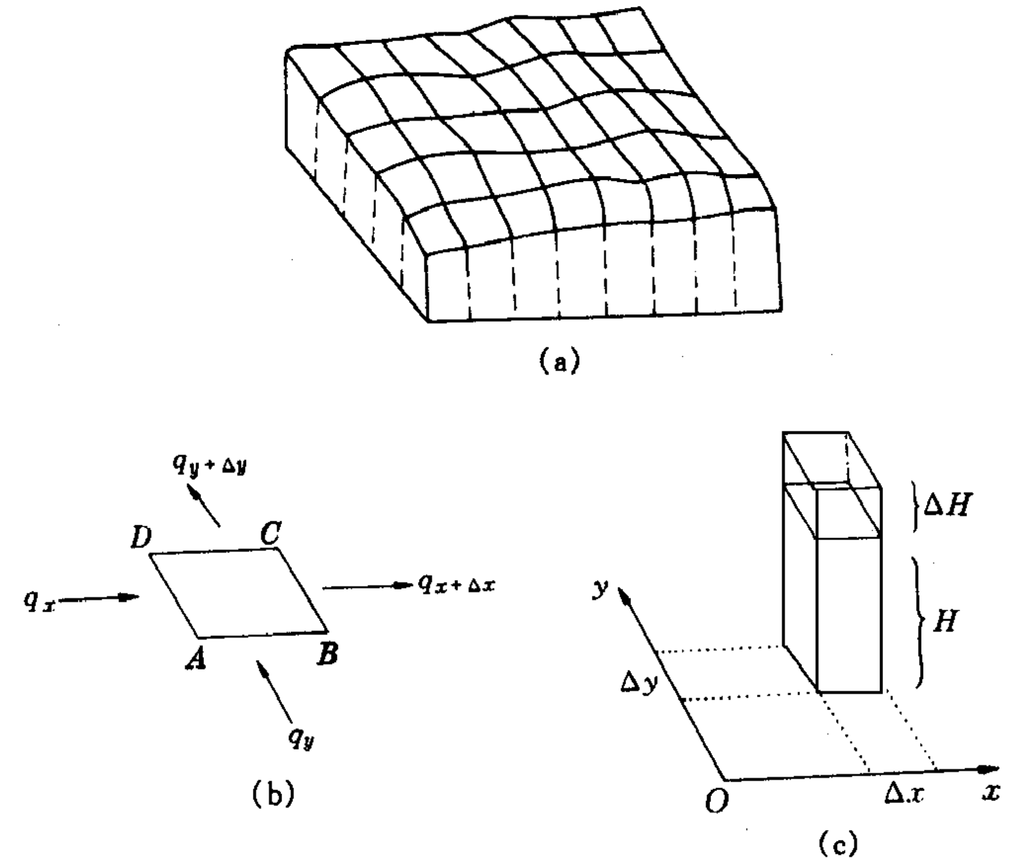


图 14-12：区域地貌离散分析

地貌是构造、过程和时间的函数，现存地貌形态的定量指标及其物质组成等特点集中反映了地貌系统以往内外营力共同作用的总和。地貌信息系统分析方法就是提取地貌系统各不同侧面的特征参数，采用系统综合分析方法取得对区域地貌发展规律的总体认识。地理信息系统方法的发展，使得建立地貌信息系统，对庞大复杂的地貌空间数据进行定量分析处理和信息复合与综合评价成为可能。

6．2 地貌信息系统

地貌信息系统基于 GIS 工具开发，集成了 GIS 的通用功能，并包括如下地貌分析模块：

1）地貌形态分析

用于提取由地表形态反映出来的地貌特征参数，包括 DEM 以及基于 DEM 的各种地貌参数：坡度、坡向、起伏度、河网密度、沟谷参数等等。

2）地貌信息复合

包括地貌参数复合模型和地貌信息/专题信息复合模型，前者将各个地貌参数复合，发现其间的规律性，得到综合的分析结果；后者考虑到地貌格局还受到外动力地质作用的影响，将地貌信息与地质、水文、气象等信息复合，发现其联系。

3）地貌区域评价

根据单个或者多个地貌特征的分析，评价地貌环境对农业、建筑、工程、旅游、居住条件的影响以及评价地貌条件对于各种自然灾害发生、发展以及防治的影响。

4）地貌成因分析

包括：

内动力成因分析模型：内动力是地壳内部能量产生的营力，决定地貌的格局与骨架，可以从几何效应和力学机制两个方面进行分析，有限元方法是研究地貌内动力成因机制的有效方法。

外动力成因分析模型：外动力是地球表面受太阳能和重力以及生物活动的营力，主要包括重力作用、风力作用、水利作用、温差作用和生物作用等，影响着地貌的具体形态。

5）地貌预测

可以采用两种模型进行地貌预测，即理论地貌预测模型和统计地貌预测模型，前者以成因分析为基础，采用数理模型进行宏观和微观的模拟，预测地貌的发展；后者在地貌发展历史数据的基础上，采用“黑箱”或“灰箱”的方法进行统计分析，预测地貌发展趋势。

7．医疗卫生

无疑，健康问题是人类生活中最为关注的问题之一，考虑到以下因素，在医疗卫生领域应用 GIS，将有助于解决健康问题以及制订相应的政策。

1）某些疾病只是限于一些地区（地方病），探讨其局部的自然和社会条件，对于研究疾病的原因以及治疗手段有着重要的意义。

2）医疗设施（医院、诊所、急救中心等）具有空间分布特性，在一个区域内合理分布医疗设施有利于资源的有效利用。

7．1 GIS 与流行病研究

由于流行病是用于描述和解释某种疾病的发病率，从空间的角度来看，流行病学需要很好地描述流行病发病率空间分布特征的手段，进而可以研究发病率模型，以发现流行病和周围环境的关系。通常，GIS 在流行病研究中主要提供了如下三个方面的功能：

7．1．1 流行病数据的可视化

假定要使用一组点来表示某个人群的发病率，而通常无法直接获得点数据，而只能获得整个区域——如人口普查区或其它行政区划单元——的数据，这些数据包括了疾病的发病率及各疾病的年龄结构。

由于各个空间区域的大小各异，使得表现发病率的空间分布时不直观，一个解决的办法是采用比较统计地图（Cartogram），在比较统计地图上，各个单元的大小与具有发病危险的人口数目成比例，然后将发病率绘制在该图上，这样可以利用 GIS，清楚地将流行病数据的空间分布可视化（图 14-13）。

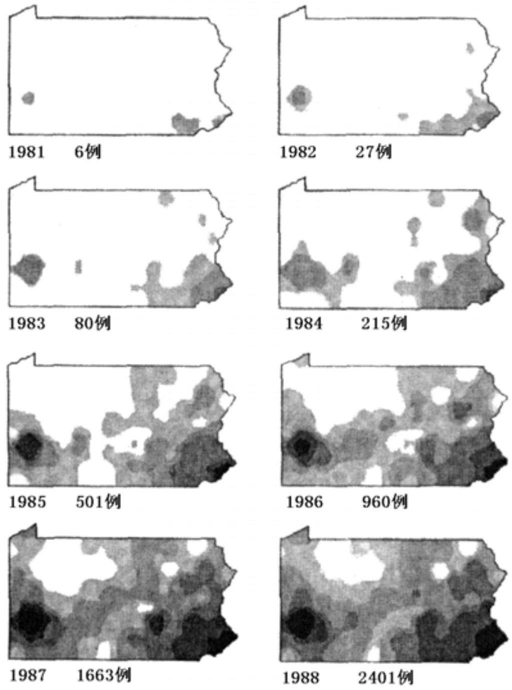


图 14-13：宾西法尼亚州 1981-1988 年爱滋病的扩散情况，
图上灰色的区域成几何级数增长生动地表现了爱滋病的扩散

7．1．2 空间数据分析

实际上，空间数据可视化与空间数据分析的界限是十分模糊的，如制作比较统计地图需要对地理空间进行变换。由于各个区域人口的分布是不均匀的，为了分析流行病病例的空间分布，通常采用的一种方法称为密度估算（Density Estimation）。该方法是采用一个移动窗口（Moving Window）覆盖于栅格化的位置点上，计算每个窗口内的密度，采用这种方法，关键是确定窗口的大小。此外，计算发病率密度与人口密度的关系；以及对病例空间分布进行聚类分析（Cluster Analysis），并探讨每一特定聚集区域具体特征，都可以应用于流行病数据的空间分析。

7．1．3 流行病模型

在识别出每个流行病的空间分布聚集区之后，一个重点的研究是确定流行病和其周围环境的关系，这将有助于验证对流行病发病原因的假定。做法之一是将空气、水质的空间分布与流行病的空间分布进行对比研究，例如在非洲疟疾传染的研究中，使用利用遥感图像计算的 NDVI 指数来进行建模。此外为了得到流行病的源地，采用距离分析，通常是根据预先假定的病源地进行缓冲区计算，然后得到每个距离范围内的病例数目，进而确定真正的病源地。利用空间分析手段确定疾病原因，最著名的案例是英国医生琼•斯诺利用地图发现霍乱病病源：

1854 年 8 月到 9 月英国伦敦霍乱病流行时，当局始终找不到发病原因，后来医生琼•斯诺博士在绘有霍乱流行地区所有道路、房屋、饮用水机井等内容的 1:6500 城区地图上（图 14-14），标出了每个霍乱病死者的住家位置，得到了霍乱病死者居住位置分布图，他分析了这张图，马上明白了霍乱病源之所在——死者住家都集中于饮用“布洛多斯托”井水的地区

及周围。根据斯诺博士的分析和请求，当局于 9 月 8 日摘下了这个水井的水泵，再往后就没有出现新的霍乱病人了。

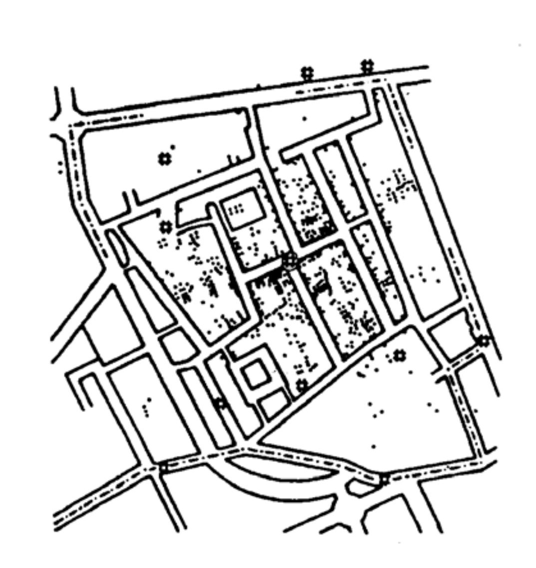


图 14-14：1954 年英国霍乱病流行地区霍乱病死者居住位置分布图

7．2 GIS 与医疗设施分布

7．2．1 医疗设施规划

不论是在发达国家还是发展中国家，随着“规划需从当地的实际情况出发”的观点的不断被认可，一种以医疗中心为主的观点也不断被采纳。这意味这不再为了所谓声望而投资大型医疗基地，而代之以建造更多面向小社区的诊所，以方便大众。这时，显然需要一种办法来提供人口统计和发病率的详细资料医疗设施需求。在进行这种医疗设施定位时，GIS 可以发挥的重要作用。图 14-15 显示了在英国 west Sussex 的研究中，通过确定病人到医生的流量，从而对该地区医疗设施规划的合理性做出判断。

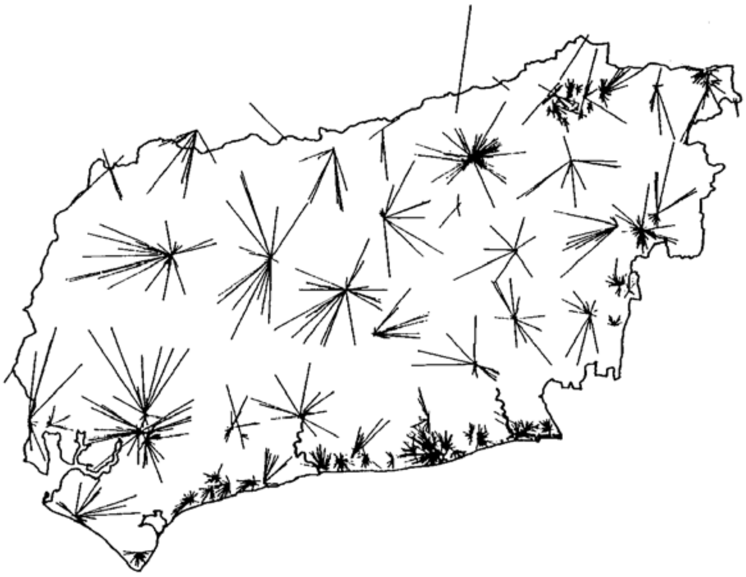


图 14-15：病人到全科医师的主要流动方向

反应了每个全科医师的“吸引范围”（west Sussex）

在定义医疗需求区域后，为了评估当地居民的需求，通常需要将已获得的调查统计数据进行社会经济分类，以便更好地描述那些小区域。这样，拥有不同生活方式的不同阶层，或不同生活方式与发病率和死亡率之间联系的数据就可获得。然后将标准化发病率通过这一阶层的情况来推及个人，从而产生综合疾病指数应用于实践。这个综合疾病指数可用于预测病人药品的开销等方面。

7．2．2 可达性，可用性及结果

GIS 在理解医疗结果与医疗设施可达性的关系方面很有帮助。对于一些疾病，如交通事故，迅速送往医院对于病人的抢救至关重要，这需要医院具有良好的可达性，对于可达性与医疗关系的探讨，同样有助于规划医疗设施。

7．3 联系流行病学与医疗实施规划—空间决策支持系统

在医学地理学中，地理流行病学和医疗设施规划是两个最主要的领域。建立两者间的桥梁，需要包括空间决策支持系统，如果分析显示在特殊的局部区域存在健康问题，此时，空间决策支持系统将提供工具及策略来处理这一问题。在研究中，发现以下问题是值得注意的：

1）地图的比例尺，小比例尺的地图会掩盖许多细部的重要问题，如小区域间发病率的变化。大比例尺地图则可使有关当局集中力量于最需要他们的区域。

2）疾病的传播是无国界的，随着人类活动的扩展，疾病也随之传播，通过对人口流制图，可以研究疾病的传播规律。

8．军事

军事是以准备和实施战争为中心的社会活动。一切军事行动都是在一定的地理环境中进行的，地理环境对军事行动有着极其重要的影响与作用。随着人类社会向信息化迅速发展，

未来高技术战争中信息对抗的含量将越来越高，特别是高技术条件下的局部战争，由于战争爆发突然，战争进程加快、战机稍纵即逝等特点，对作战指挥的时效性有了更高的要求。指挥决策智能化、作战指挥自动化、武器装备信息化成为未来战争取胜的关键。在这种需求下，出现了数字化战场，数字化的地理环境信息已成为指挥决策的必要条件之一。因此，作为空间军事信息保障的军事地理信息系统已成为现代化军事斗争的一项重要内容。

军事地理信息系统 MGIS（Military Geographic Information System）是地理信息系统技术在军事方面的应用，是指在计算机软硬件的支持下，对军事地形、资源与环境等空间信息进行采集、存储、检索、分析、显示和输出的技术系统。它在军事地理信息保障和指挥决策中起着重要的作用。

军事地理信息系统和遥感、全球定位系统关系密切，同时和指挥自动化系统 C^3I（Command，指挥；Control，控制；Communication，通信；Information，情报）紧密地联系在一起，形成一个多功能的统一系统。它一般由六个子系统组成：信息收集子系统、信息传递子系统、信息处理子系统、信息显示子系统、决策监控子系统和执行子系统。其中，情报是军事决策的基础；信息收集、处理和显示是系统的核心；通信和控制是信息传输和决策过程的保证；指挥使军事决策具体执行。地理信息系统技术在情报的收集、处理、显示和指挥决策方面发挥着重要的作用。

另外，由军事技术革命引发的数字化战场建设已成为未来战场发展的主流，建设数字化战场和数字化部队已成为 21 世纪军队发展的大趋势，引起了各国的普遍关注。美国著名未来学家托夫勒指出，建设数字化战场是一项比研制原子弹的“曼哈顿工程”更具挑战性的系统工程，“数字化战场是打赢信息战的关键”。战场数字化就其内容来讲，主要是战场地理环境的数字化、作战部队的数字化、各种武器的数字化和士兵装备的数字化。从某种意义上来讲，战场地理环境的数字化是其他数字化的基础，它为作战部队和各种武器装备的数字化提供了必需的战场背景环境和空间定位基础。

8．1 国外军事地理信息系统（MGIS）现状

军事地理信息系统在海湾战争及以后的战争中发挥了重要的作用，受到了各国军方的普遍重视。世界上大部分国家都建立了用途不同、规模大小不等的军事地理信息系统。报道较多的是美国、俄罗斯、英国、澳大利亚等国的 MGIS。其应用领域包括：

1）基础地理信息 包括地形图、DEM、DTM 等。

2）航海、航空管理 航海图、制定计划航线、障碍物、禁区、助航设施、导航管理、空中交通控制等。

3）地形分析 包括战场模拟、行军路线、应急线路分析、越野机动、涉水分析、通视点分析、距离量测、面积量测、武器打击轨迹分析等。

4）任务规划（战略层次） 包括军事基地规划、军事基础设施管理、打击效果评估、巡航导弹支持、战区规划、入侵应急规划、目标分析、轨道建模等。

5）战争管理（战术层次） 包括战场监测、战场管理、小战区规划、登陆计划、战术模拟、后勤保障规划、交通规划等。

6）基础作业支持 包括拦截应用、环境应用、军事设施分类规划等。

7）边界控制 包括边界巡逻和交叉分析、毒品禁运、移民控制等。

8）情报 包括反毒品活动、反恐怖主义活动、武器监视与跟踪、情报收集等。

8．2 国外军事应用系统示例

美国国防制图局（DMA，Defense Mapping Agency）是美国军事 GIS 的管理部门，1989 年以来投入大量的人力和物力来开发军事 GIS 系统。例如：

1）数字航海图 数字化海图（Digital Navigation Chart，DNC）和 GPS 结合用于海军的导航，来取代以前使用的 4500 张纸质海图。这些数字地图可以支持广泛的空间分析和地理查询。数字化海图一般包含有 12 个特征数据库：地名注记、水文要素、界限、港口设施、水下地形、航道、航海标志、救援标志、海洋环境、陆地地形地物、航海障碍物、数据的优先级。

2）数字地图 VMAP（Vector smart MAP）数据分两个层次，分别对应于 1：25 万比例尺和 1：5 万比例尺的地形图。该系统包括有 10 个数据层：交通、网络、界限、水文、城市、公用设施、植被、等高线、工业设施。

3）水文资源评估系统 用于水文和测深数据的收集。

实际上，由于军事保密的特殊性，我们对国外许多功能强大的 MGIS 了解得都很少。作为军事情报的一个设施，美国正在建立和完善军事应用的全球空间信息和服务（Global Geospatial Information and Services，GGI&S）系统。其中的数据包含两个层次，国家性的和全球性的，符合 NOTA DIGEST、ISO TC/211 以及 OGIS 标准。

另外，英国国家遥感中心（British National Remote Sensing Center， NRSC）应用法国的 SPOT 图像与 MGIS 结合，模拟敌方的三维地形，对军方飞行员进行模拟训练，取得了较好的效果。

海湾战争的实践证明，利用 MGIS 技术和遥感技术相结合，可为战时提供实时的地理信息保障。美国国防制图局（DMA）在海湾战争的战场上的 MGIS 实时服务，主要包括利用自动影像匹配和自动目标识别技术处理卫星和高低空侦察机实时获得的数字影像，及时地为军事决策提供 24 小时的实时服务。对于冲突地区，最基本的地理信息保障是为作战部队提供作战地图。与传统的地图制图技术相比，MGIS 的制图功能无疑是出色的，它既可以提供数字地图，也可以输出精美的印刷地图，其生产速度比传统的制图方法要快得多，特别是在遥感技术支持下，可实时地获得制图数据。在海湾战争期间，美国国防制图局进入一天 24 小时生产状态，共开发了 12000 套新的地图产品，其中 600 套数字地图，印刷了 100 万幅战地地图，这些地图覆盖科威特、沙特阿拉伯、伊拉克、叙利亚等国和地区，比例尺为 1：5 万不等。由于对这些战地地图的迫切需要，多国部队在这次战争中使用 C130 远程运输机运送军事地图到沙特阿拉伯，优先于药品的运输，仅次于爱国者导弹发射架部件的运输，由此可见，战时的实时地理信息保障是何等重要。

除上述地图产品外，美国国防制图局和工程地形实验室还使用了数字地形高程数据、栅格图形数据、地名字典等。在没有现成地图资料的地区，利用 MGIS 处理 SPOT 和 LANDSAT 影像及机载雷达影像，得到可直接应用的影像地图。因此，MGIS 和 RS 相结合，可较好地完成战时的实时地理信息保障任务。

8．3 MGIS 的发展方向

概括而言，军事地理信息系统有以下的发展趋势，以便更能够满足现代战争的要求：

1）基础数据、基础设施大型化。

2）各领域应用专业化。

3）实用化，简单化，易学易用。

4）小型化。

5）标准化。

三个法国士兵打不过两个马穆鲁克人，100 个法国士兵可以与 100 个马穆鲁克人站成平手，200 个法国士兵则绝对可以战胜 300 个马穆鲁克人.

拿破仑

第十五章 地理信息系统应用项目组织和管理

导读：本章首先简单概括了 GIS 应用项目，提出了 GIS 应用的三元划分，然后介绍了几种 GIS 应用的建立途径以及项目周期。在项目开始之前，需要进行策略性规划和实施规划，如果要进行项目承包，则要签订合同;在 GIS 项目进行过程中，要进行人员管理，数据管理以及质量控制，本章根据 GIS 具体技术特点，提出了相应的建议。在最后，介绍了软件研制和开发的质量控制的两个标准，ISO9000 系列标准和 CMM 模型，作为项目开发机构的指导。

本章和地理信息系统软件工程技术一章讲述的内容是互相依赖的，后者重点在于技术，而本章则侧重于组织和管理，往往组织管理对项目成败的影响比所采用的技术更大。

1．GIS 应用项目简介

1．1GIS 应用模式与分类

地理信息系统可以区分为通用地理信息系统平台和面向特定专题或地区的 GIS 应用，前者提供了基本的空间信息处理方案，可以应用于各个领域，一般由专门的软件开发商完成，其开发过程类同于普通软件系统（如数据库管理系统、桌面出版系统等等）的开发；而后者往往针对某个用户单位提供特定的技术手段。具体而言，GIS 应用按照其应用模式又可以分为两类，即科学研究工具和办公服务系统两种。

1）科学研究工具

将 GIS 作为科学研究工具的应用模式，强调对于科学计算结果的获得和分析，把 GIS 作为科学研究的辅助手段。它主要应用于有关地学领域的科研项目研究中。它不仅需要用到地理信息系统通用软件所提供的功能，而且还要用到各种专业分析模型。

2）办公服务系统

办公服务系统应用于涉及空间数据的政府部门以及企业，以提高管理效率、制定好的决策和实现组织目标。

办公服务系统按照其应用层次的高低，又可以分为空间事务处理系统(STPS-Spatial Transaction Process System)，空间信息管理系统(SMIS-Spatial Management Information System)，空间决策支持系统(SDSS-Spatial Decision Support System)和专家系统（Expert System)。

事务处理系统的目标是迅速、及时、准确地处理大量空间信息，能够有效地进行日常事务的自动化处理。它注重于空间数据的收集，处理和存储，以供将来使用在各种大型应用地理信息系统的数据采集部门和具体事务部门都有着广泛的应用，包括测绘、资源调查、地籍管理、地图出版等领域。

空间管理信息系统是基于空间事务处理系统发展起来的，除了提供高效率的信息处理以

外，还对决策者提供辅助决策信息，包括数据的查询和统计以及专业模型的分析功能。SMIS 运用专业模型来处理和分析数据，以实现对业务工作中确定性问题的处理和管理，提供决策服务。

SDSS 为决策者提供了一个模拟决策过程，并提供了选择方案的决策支持环境，强调系统推理的有效性，更多地应用于宏观决策过程。

专家系统是能够模仿人工决策处理过程的基于计算机的信息系统，它由知识库、推理机、解释系统、用户接口和知识获得系统组成。它扩大了计算机的应用范围，使其从传统的资料处理领域发展到智能推理上来。SMIS 能够提供信息帮助制定决策，SDSS 帮助能够改善决策质量，只有专家系统能够应用智能推理制作决策并解释决策理由。

上述应用模式之间的界限并不是绝对的，一个决策支持系统可以使用与科研工具中一致的分析模型，并且广义上讲，所有的地理信息系统应用的最终目标都是为了进行空间决策。

除了从应用模式上划分 GIS 应用，从规模上划分为小型、中型和大型应用。小型 GIS 使用数据量小，使用系统的用户少，主要针对一个部门或特定领域，注重于专业模型的开发和应用。中型 GIS 应用于多个部门，数据量大，运行于局域网或城域网环境，侧重于决策支持。大型的 GIS 应用则拥有非常多的用户和海量的数据，注重数据的管理，并通过网络实现分布式的计算和数据管理，并通过 Internet 发布空间信息。

另外，再考虑到 GIS 具体应用的领域，可以对 GIS 应用进行三元划分（图 15-1），每个 GIS 应用都对应于该三维空间特定的坐标（表 15-1）。

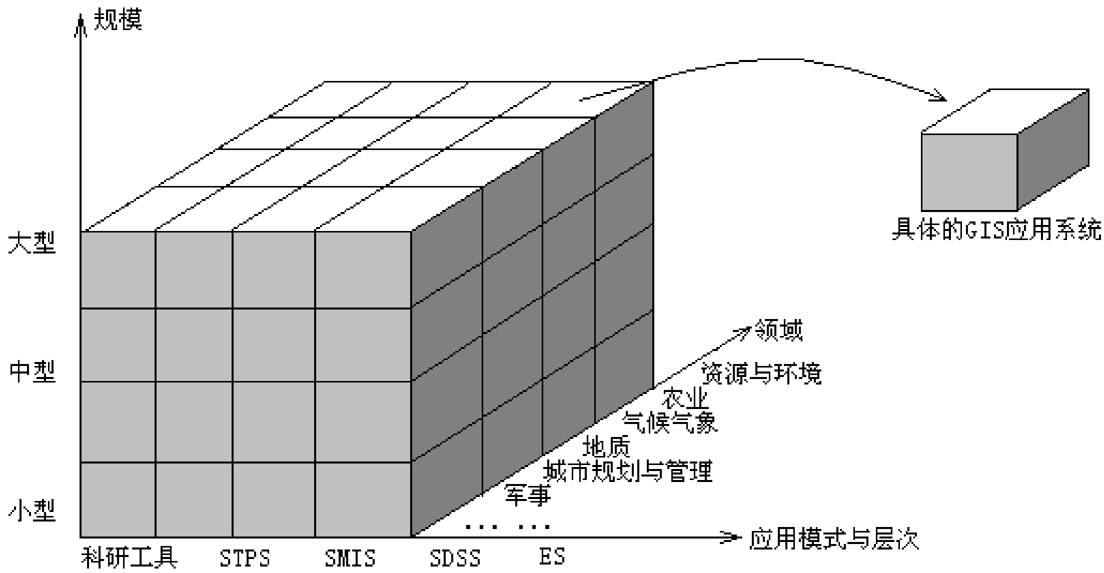


图 15-1：地理信息系统应用的划分

表 15-1：GIS 应用的三元划分示例

名称	规模	应用模式和层次	领域
XX 小流域水土流失系统	小型	科学研究工具	环境
XX 城区土地划拨系统	中型	空间事务处理系统	地籍
XX 林场管理系统	中型	空间管理信息系统	林业
XX 大城市市政管理系统	大型	空间管理信息系统	城市管理
XX 省可持续发展决策支持系统	大型	空间决策支持系统	资源与环境

从应用领域的角度对 GIS 的划分

GIS 的数据模型可以划分为场模型和要素模型，图 15-2 表示了这两类模型在不同领域的应用。



图 15-2：根据场模型和要素模型对 GIS 应用领域进行划分[N.R. Adam& A Gangopadhyay]

1．2 开发 GIS 应用的方式

一旦组织决定建立用于本单位的地理信息系统，可以采取三种方式，一是通过购买，在组织中实施；二是请软件开发商来开发；三是由组织内部的人员开发，可以是专业的地理信息系统人员，也可以是最终用户。

其中购买 GIS 又可以分为购买 GIS 平台软件进行二次开发，购买完整的软件产品，购买完整的 GIS 系统以及购买 GIS 服务。

请软件开发商进行开发以得到完整的软件产品或系统，也可以认为是系统购买的一种形式，但是从购买到系统运行时间比直接购买时间要长。在请开发商进行开发时，可以有两种具体方式，即承包开发和合作开发。前者由开发商独立开发完整的系统，又称为“交钥匙”系统；后者在开发过程中，可以由用户方的技术人员参与开发，双方合作完成整个系统，这种方式有利于开发系统时对问题域的准确把握。

上述的 GIS 应用建立方式各有利弊（表 15-2），各个组织可以根据具体情况确定采用何种方案。

表 15-2：建立 GIS 应用方案的比较[据 Aronoff 扩展]

实施方案	用户开发	购买通用平台	购买完整软件	购买完整系统	购买服务	承包开发	合作开发
对提供者依赖性	低	低	高	很高	很高	很高	中
到系统运行时间	长	长-中长	短	很短	很短	长-中长	长-中长
初始费用	低	中等	中等	高	高	高	中等
人力费用	高	中等	低	低	很低	低	中等
风险和不确定性	高	较低	低	低	中等	高	中等

灵活性	完全可以	完全可以	中等	中等	不定	高	完全可以
对用户技术要求	很高	高	中等	中等	很低	中等	高
现有资源的利用	高	高	中等	低	很低	低	中等-高

1．3 GIS 应用项目

如果要建立 GIS 应用的组织不是购买现成的、直接可以使用的 GIS 系统，那么就意味着要进行系统开发工作，无论系统是自行开发，还是承包开发、合作开发。从开始确立需求，到最终的系统投入运行，这一个时间段内的活动统称为 GIS 应用项目。

通常一个项目包括四个阶段，识别需求，提出解决方案，执行项目以及结束项目，这四个阶段构成了项目的生命周期（图 15-3）。

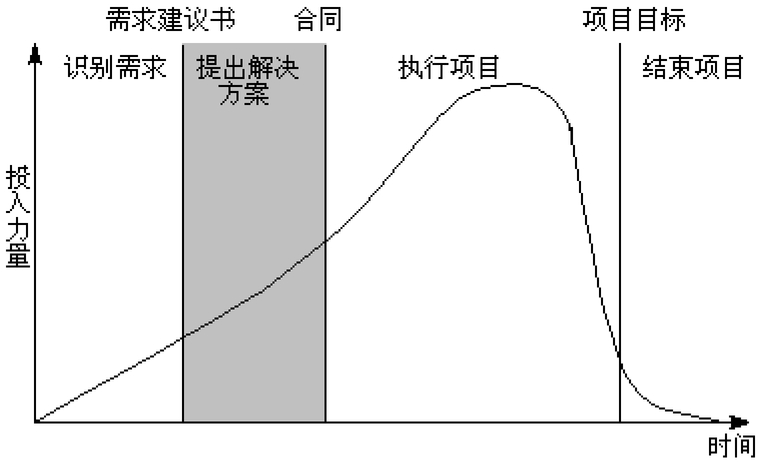


图 15-3：项目生存周期[J. Gido]

识别需求由准备建立 GIS 应用系统的组织进行，包括认识当前组织的存在问题，准备用 GIS 解决哪个方面的问题，分析项目可行性以便进行项目预算等方面的问题。

如果是用户方自行开发建立 GIS 应用，那么提出解决方案、执行项目以及结束项目都是由用户方来完成。如果需要由 GIS 开发商进行开发，那么需要向他们征询解决方案，并签订合同，前者可以通过招标的方式进行。在签订项目合同之后，项目执行主要由开发方负责，用户方配合开发方的工作，直至整个项目结束。图 15-4 描述了这两种地理信息系统建立过程。

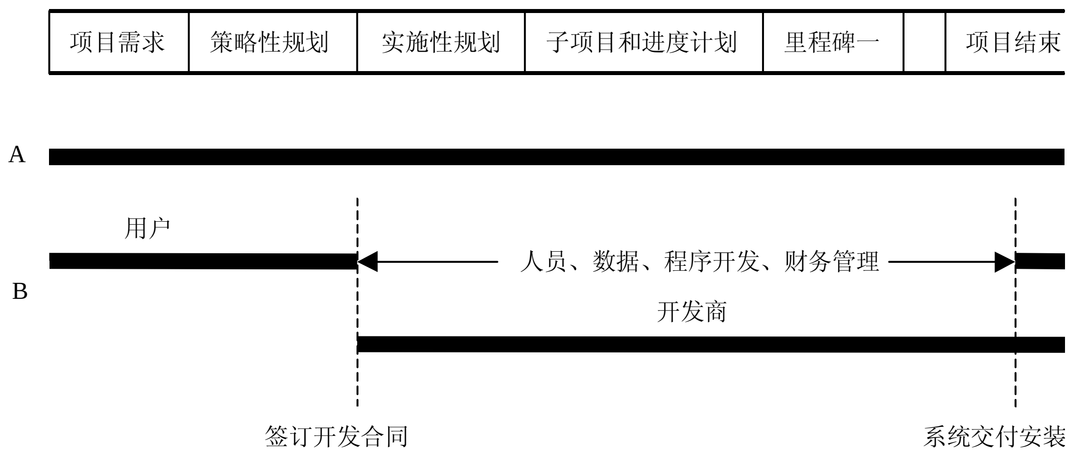


图 15-4：地理信息系统项目过程
（A：表示用户自行建立系统；B 表示通过招标由 GIS 开发商建立）

2．应用项目策略性规划

越来越多的机构都在开发 GIS，但是根据调查，有大量的 GIS 系统不能真正地完成并正常运行，造成这些项目失败的大都不是由于技术上的问题，而是组织和管理中的问题，而不能进行细致的项目规划是根本的原因。项目规划包括策略性规划和实施性规划，前者在项目正式启动之前进行，而后者则往往是项目实施的第一步工作。

造成地理信息系统开发失败的因素[村井俊治]，从相反的角度说明了如何才能开发一个成功的地理信息系统应用。

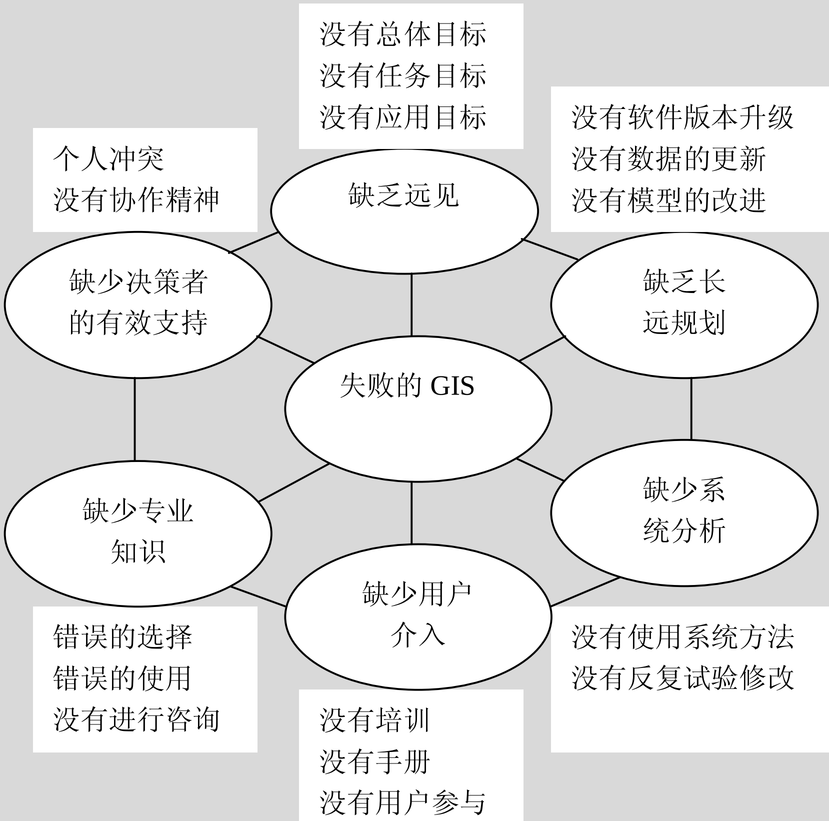


图 15-5：造成地理信息系统开发失败的因素

应用项目策略性规划工作对应于项目生命周期的识别需求阶段，主要目的是确定要建立怎样的 GIS 应用，组织现有的条件，包括资金、人员、设备、场地等等，能否支持 GIS 的建立和应用。如果用户自行开发 GIS 系统，可以包括系统开发方案的选择；反之，一般由开发方提出解决方案。

GIS 项目策略性规划一般包括五个主要部分：

1）机构的状况分析和调查

从 GIS 的角度了解一个机构目前的组织、任务和技术状况，包括责任范围，发展目标，技术力量，对使用信息技术和 GIS 的态度，人力、物力和资金状况等等。状况分析和调查为下一步计划提供了概念化的基础。

2）机构的 GIS 系统开发目标

明确定义整个机构 GIS 总的发展方向，这个方向应当与整个机构的任务和责任范围一致，并且应当对整个系统各方面的能力作出明确的定义。制定总方向的过程需要机构内高程

管理人员的参与，他们对于整个系统开发的必要性的理解和认识会对整个系统的发展有所帮助。

3）系统开发的可行性

主要包括资金、技术力量和组织运作三个方面。资金的分析应当考虑到整个 GIS 实施、运行和维护的全过程，通用的方法是成本效益分析。在技术可行性方面，要考虑 GIS 项目中所要求的技术能否满足，技术发展以及新技术出现对项目的影响，是否需要对人员进行技术培训。组织方面，包括整个机构能否愿意承受引入 GIS 技术所带来的变化以及能否在开发过程中相互协作完成开发任务。

在进行可行性分析时，不可忽视各个方面的变化所引发的风险，要对风险进行客观的评价，并作出相应的防范措施。

4）系统开发的方案选择

系统开发的方案选择覆盖了 GIS 系统的各个方面，包括数据部分、功能部分、系统的整个软件和硬件的框架等，以确定各个方面的实施部门、进度安排和管理方法等等。

5）GIS 系统决策规划书的起草

方案确定之后，需要将方案起草成文件形式，由有关部门审批，同时也可以作为系统规划和实施的指导性文件。

在策略性规划完成并得到批准之后，如果是由用户自行开发系统，则可以直接进行下面的开发工作；如果需要由另外的 GIS 开发商合作开发或承包开发，则可以以需求建议书（RFP, Request for Proposal）的形式提交给开发商，而开发商将其作为进行需求分析的重要依据，此外，在合同中也体现了策略性规划的部分成果。

3. 应用项目合同

如果用户单位确定由其它软件开发商承包或合作开发 GIS 应用软件系统以及向开发商购买现成的软件系统，则需要由双方签订关于本次开发或购买——统称为软件贸易——的技术合同。技术合同确定了贸易的目的，双方的权利和责任等多个方面的内容。在中国，技术合同的签订、执行、变更与终止需要在《著作权法》和《技术合同法》中的有关原则的规范下进行。

由于当事各方的具体情况不同，合同的具体内容和形式也不同。一般而言，采用双方合作开发的形式，因为用户拥有部分版权，合同中对于用户限制较小；而采取“交钥匙”的开发方式或者直接购买软件，则对作为接受方的用户[*]的限制较为严格。具体地，软件技术合同通常所包含的条款有：

1）引言

是合同的第一个条款，包括合同签订日期，当事双方名称、地址，合同背景。

2）定义

在定义部分，对一些关键性的用词进行定义，避免双方对合同中用词和概念的不一致带来纠纷。对于信息系统技术合同，需要进行定义的名词包括“被许可软件”，“文档”等等。对于“被许可软件”，需要指明软件名称、版本、形式（目标代码，源代码）；而对于“文档”，要说明它是用户手册，还是工程文档，或者两者都包括。

3）标的

说明了开发系统名称、合作开发方式、使用许可以及系统的技术规格。技术规格又包括

---

[*] 在本节中，“用户”并不仅指单纯的系统使用者，也包括为了赢利而销售应用系统的单位。所以一般称为接受方，而软件开发单位则称为供应方。

软硬件环境，功能描述和性能说明。

4）授权及限制

该部分规定软件供应方和接受方的权利以及对这些权利的限制，在每一项具体的合同中，用户方获得的授权内容和限制是不一致的。在技术合同中，授权包括：

（4．1）使用权：软件使用权是通过许可合同授予用户最主要的权利，是签订许可合同的标的，不过作为供应方可以对使用权增加一些限制，包括运行环境，数据等等。

（4．2）分销权：如果接受方不是最终使用用户，可以经销系统，则需要在合同中确定其分销权。

（4．3）独占性：使用权和分销权可以是独占的，也可以是非独占的。分销权独占性一般只是在一个地区范围内和行业范围内成立。

（4．4）复制权：对用户制作合法拷贝备份的规定。

（4．5）修改权：一般来说，获得使用权或分销权的用户都无权修改供应方提供的软件。不过，根据实际的情况，可以在合同中作出更为合理的规定。

对于专业 GIS 应用软件，由于一些专业模型需要由用户提供，并且需要修改，一般规定用户有修改权，以改进模型，但在这时要对供应方的维护责任作出规定。

（4．6）对许可合同的转让权：被许可人可以将许可合同赋予自己的全部权利转让给第三方，但是只能转让一次，并且需要征得供应方的同意。

（4．7）对计算机系统的限制：通过限制软件运行的计算机系统防止系统扩散。

（4．8）地区限制：限制合同的运用地区。

5）交付与安装

明确规定供应方提供软件——包括程序和文档——的具体内容，对于程序，需要说明是否包括源代码；文档要说明包括哪些具体文档；同时对于是否由供应方将应用软件安装在计算机内。

6）移交与验收

为了保证接受方得到的系统是可以实际使用的，在合同中必须对软件的移交和验收作出安排。

7）培训

如果需要供应方为接受方承担培训工作，则应在合同中设置培训条款，包括培训时间、内容和费用承担等。

8）维护

在任何软件技术合同中，都要包括维护条款，维护包括纠错性维护、适应性维护和完善性维护，合同中首先要注意的是纠错性维护。维护条款包括修正错误的时间，费用承担，同时规定哪些情况不在供应方维护范围内。

9）期限

期限规定了合同的有效时间。

10）报酬与支付

在报酬和支付部分包括三个方面的问题：支付方式、费用支付针对哪些软件产品和服务、支付时间。例如，对于最终用户，采取的支付方式包括依次性固定支付，周期性固定支付和根据运行情况进行支付三种方式。

11）纳税

该部分规定的各项税款的承担单位。

12）所有权的归属

一般而言，接受方只是获得了软件的使用权，所有权并没有转让。但是在合作开发中，可以根据具体情况规定所有权归属。另外，在承包开发的情况下，为了保护接受方权益，可

以规定所有权属于制定的委托方。

13）保密与限制竞争

在承包开发过程中，双方都有可能使用对方的机密，需要规定保密责任。

14）担保和赔偿

担保是指供应方对自己所供应软件的担保，包括功能、性能以及版权等方面。如果所做保证不能实现，则要向接受方提供损害赔偿。

15）违约处理

规定了合同双方发生违约时的处理，包括终止合同以及赔偿等等。

16）争议处理

在合同执行中出现争议的处理，通常的做法是提请拥有计算机技术方面的专家仲裁裁决。

17）终止

如果合同没有确定期限，规定合同在哪些情况下终止以及合同终止后双方要做的事情和继续承担的责任。

18）条款的完整性

强调条款的完整性，未在合同中的内容，一律不能成为本次软件贸易的组成部分。

19）不可抗力

由于出现不可抗拒因素而影响合同的执行，通常一方应给予谅解并适当延长履行合同期限。

20）法律的选择

如果双方属于不同的国家，要说明发生纠纷时采用哪个国家的法律。

21）生效

一般合同签字后即可生效，在有些情况下，需要由有关部门批准后方可生效，在本条款中加以说明。

上面是软件技术合同的一些基本条款，除此之外，还往往包括一些附件以详细说明具体的技术问题。

4．应用项目实施性规划——软硬件配置

应用项目开始之后，就要进行实施性规划，在这一阶段，主要要确定系统软硬件设计。如果不是完全自行开发系统，软件设计过程中要进行软件配置，反之要进行软件的总体设计。

在进行软硬件配置时，主要考虑的因素有：

1）系统数据量；

2）用户数目，并发访问情况；

3）现有网络环境；

4）投资规模；

5）GIS 应用在整个系统中的位置，是一个纯粹的 GIS 系统还是与其它系统结合而成的综合信息系统。

这些因素，在策略性规划阶段就已经基本确定，此时要根据这些因素进一步选择软件以及硬件设备。在进行软硬件配置时，除了上述因素外，可以参照以下原则：

1）不要只重视硬件，而忽视软件的选择；

2）保护已有的投资，包括软件、硬件设备、数据等；

3）保证系统的可扩充性；

4）考虑技术的发展方向；

5）如果要进行二次开发，要着重考虑软件的二次开发支持能力；

6）重点考虑软硬件功能能否满足项目需求；

7）性能也是不可忽视的重要因素。

8）选择软硬件的功能必须要满足系统要求，而性能指标可以事先界定，通常系统的性能与投资多少直接相关，这就需要根据投资数额确定一个最佳性能/投资比，不至于引起浪费或者不能满足项目要求。

下面给出了进行 GIS 软硬件配置时的一些具体的意见（表 15-3）。

表 15-3：GIS 软硬件配置时的一些具体的参考意见

项目	具体考虑因素
软件	
数据量	如果数据量大，要关注软件对大数据量处理的性能，是否支持图库
数据的分布性	如果数据是分布管理的，需要软件支持网络
已有数据	软件能否顺利地读入现有的数据
需要二次开发	软件是否具支持二次开发，支持方式（API，控件，宏语言）
系统需要与其它软件协作	软件对数据库是否支持
现有的软硬件平台	软件必须能够运行于现有的软硬件平台上
需要与 Internet 结合	软件是否支持 Web，具体方式(ActiveX,Java 等)
硬件	
数据量大，并且需要共享	选择数据服务器
数据需要定期备份	磁带机，光盘课录机
大量的数据录入	扫描仪，数字化仪
地图输出	彩色喷墨打印机
高质量、大幅面地图输出	喷墨绘图仪，其它高档绘图仪
数据采集	GPS 接收机，平板测图仪
移动计算	笔记本电脑
已选择的软件	使软件能够运行良好的硬件平台
网络	
局域网	细缆线/双绞线，集线器
远程访问	拨号网络支持
接入 Internet	各种专线

5．子项目划分和进度安排

在一个较大规模的地理信息系统建设过程中，由于要涉及各个部门，功能和数据数据的形式多种多样，并且要涉及不同的软硬件平台，需要在项目实施之初，制定计划以协调各个方面的问题[*]。通常，计划中最重要的内容是划分子项目以及整个项目进度的安排。

5．1 划分子项目

子项目的确定是整个 GIS 系统设计过程中最重要的环节之一，子任务的合理制定使得各方面的细节不至于遗漏，使整个计划、预算和质量得到保证。子项目不仅要包括技术任务，

[*] 到项目开始执行，不再区分用户方和开发方，原因是无论由哪一方进行开发，进行的项目组织管理工作都是相似的。

而且要包括各种非技术的任务，包括管理、支持等等。划分子项目时，可以按照整个项目、项目、任务、子项目、具体工作、步骤等六个层次，进行自顶向下地划分。表 15-4 列出了常用的 GIS 项目任务。

表 15-4：常用的 GIS 项目任务[宫鹏]

项目规划	GIS 分析
1. 可行性分析	1. 分析模型定义
2. 市场调查	2. 分析模型的过程设计
3. 技术调查和评价	3. 分析模型实施
4. 起草报告	4. 分析制图
试点项目	5. 分析制表
1. 数据收集	6. 分析自动化编程
2. 数据数字化	7. 分析报告生成
3. 数据转换	应用系统开发
4. 数据质量控制	1. 系统的用户需求分析
5. 制图	2. 系统的设计
6. 设备购买	3. 系统设计报告起草
7. 设备安装	4. 系统编程
数据库生成	5. 系统测试
1. 数据库概念设计	6. 系统运行报告和安装
2. 数据库详细设计	7. 系统培训
3. 数据收集	8. 系统的用户报告
4. 数据数字化	9. 系统维护报告
5. 数据转换	10. 系统维护的技术服务
6. 数据编辑	其它
7. 数据质量控制	1. 人员技术培训
8. 数据修改	2. 项目管理
9. 自动化编程	3. 系统维护
数据输出	4. 数据安全备案
1. 制图	5. 项目技术会议
2. 数据制表	6. 项目中期报告
3. 自动化编程	7. 项目终期报告

5．2 进度安排的原则和表示方法

项目进度安排是管理者在进入设计和实施阶段之前需要完成的，要在时间和顺序上安排各个子项目。在进行进度安排之前，首先必须估计每项活动从开始到完成所需要的时间，其次要考虑的因素包括活动之间的依赖关系（必须完成一项才能进行下一项）以及各个活动的最早开始——结束时间和最迟开始——结束时间（例如：整个项目工期为 120 天，某项活动需要 30 天，那么它的最迟开始时间是第 90 天）。计划要有灵活性，可以根据变化进行相应的调整；此外，要保证参与人员有足够时间来完成各项任务，在任务之间安排一定的“机动时间”是一个较现实的办法。

表示项目进度的常用方法有里程碑表示法(Milestone Chart)，甘特图法（Gantt Chart），关键路径法（CPM-Critical Path Method）和墙纸法（Wall Paper Method）。

5．2．1 里程碑表示法

里程碑表示法适用于较小规模的项目，可用于起草项目计划书，该方法将主要任务作为一个阶段来处理，清晰并且容易使用。缺点是不能表达出各项任务之间的关系，对项目进程的控制能力差。表 5 给出了该方法的样本。

表 15-5：里程碑表示法

任务编码	主要内容	负责小组	预计完成日期	实际完成日期

5．2．2 甘特图法

甘特图法又称直方图法、柱状图，是目前使用较多方法，它容易表现活动的时间长短和先后顺序，但是不能描述影响关系，难以进行复杂项目的控制。图 15-6 是甘特图法的一个例子。

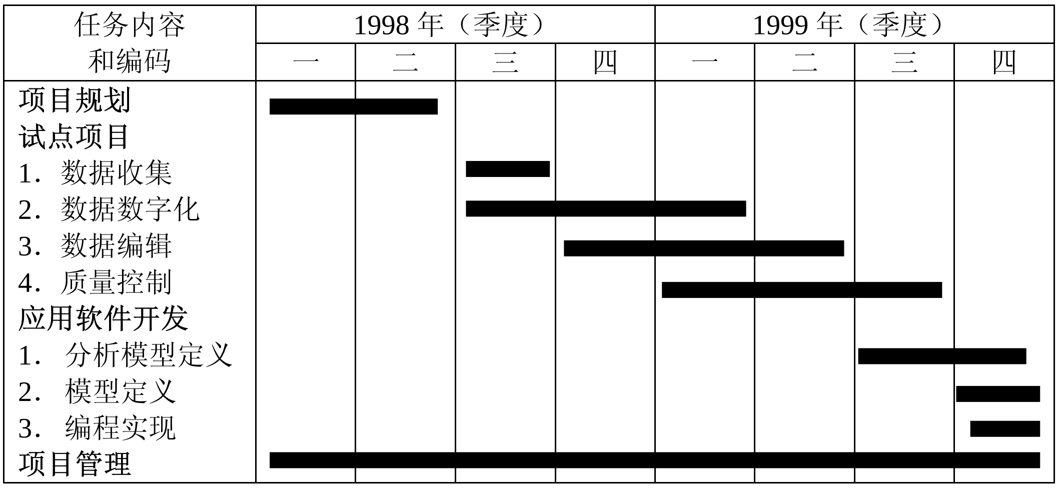


图 15-6：甘特图法表现项目进度

5．2．3 关键路径法

关键路径法表达了项目中各个任务之间的先后顺序和制约的双重关系，适用于对大型的复杂项目进行进度控制。采用关键路径法需要管理人员对子项目之间的关系十分清楚，其缺点是不够灵活，并且制作烦琐。

5．2．4 墙纸法

墙纸法主要适用于有较多人员参加的项目类型，它的特点是直接对每个具体人员安排任务。采用墙纸法制定项目进度时，需要一开始就有有关项目人员的参加，便于发挥其主动性。

在进度安排时，可以利用计算机软件辅助进行，使进度更加合理。

6．项目预算

6．1 项目预算的主要内容

项目预算是起草项目计划书时必须做的一个步骤，也是 GIS 项目管理人员最常做的事情之一。一般来说，GIS 项目的支出主要有七类，与某一 GIS 项目直接发生关系的费用称为直接项目费用，否则被称为非直接项目费用。

1）人力费用：指拥有和使用人力资源来完成任务的费用，主要包括员工的工资和福利，人力费用属于直接项目费用。

2）材料和办公用品费用：包括购买各种办公用品、书籍、图件等的费用。

3）设备费用：主要包括计算机硬件的购买或租赁费用。

4）软件费用：计算机软件的购买或租赁费用。

5）软件和硬件维护费用。

6）内部服务费用：主要包括机构内部某些部门的维持费用，它们间接支持项目和整个机构的运行。此外，还包括保险、办公场所租赁、通讯费用等等。

7）其他费用：不属于以上任何一类的费用，包括旅行费用、培训等。

上面除人力费用属于直接费用以外，其他都属于非直接费用，在计算时可以用人力费用乘以一个权值以估算非直接费用。

6．2 项目预算的方法

在进行 GIS 项目预算时，要综合考虑各种费用，进行预算的方法主要有四种：上溯法，下溯法，单价法和根据项目参加人员的费用做预算的方法。

6．2．1 上溯法

上溯法适合于根据项目任务确定来确定最终的总报价的情况。图 15-7 表示了利用上溯法进行预算的步骤。

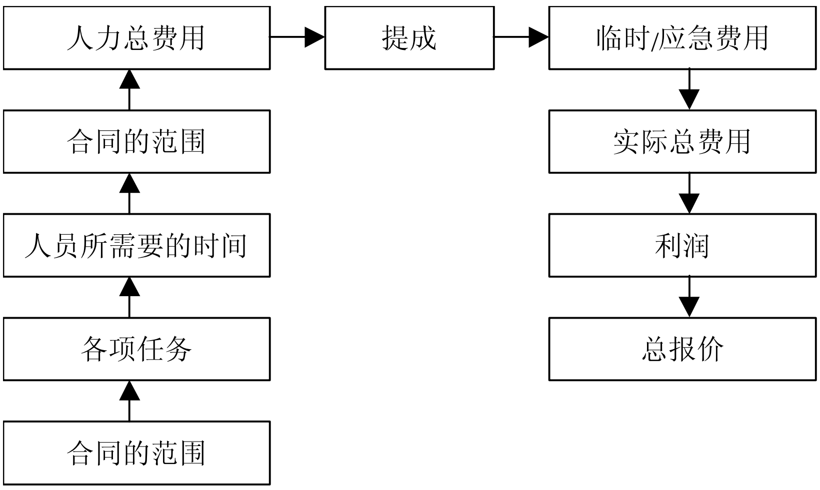


图 15-7：利用上溯法进行项目预算

利用上溯法进行预算，需要注意以下问题：

1）需要完全理解项目的范围，因为该方法是依据项目的内容和范围来着手和定义，确定范围过大或过小，都会对总报价有很大的影响。如果范围边界存在模糊，可以使用一些约束条件。

2）项目任务的划分，如果一个项目的任务划分基本正确，那么总报价不会出现太大的偏差。在进行任务划分以及任务工作量估计时，可以参照以往类似的工作进行。

3）提成和利润，在总报价中，提成和利润占有相当的份额。在美国，GIS 项目中，人力费用、办公费用和公司利润大约各占三分之一。

6．2．2 下溯法

下溯法适用于预算总额已经确定的情况，对各个划分的任务确定其具体费用。在实际工作中，可以与上溯法相结合，对预算方案进行调整，使结果既能考虑到项目总费用，又能顾及具体的任务工作量。

6．2．3 单价法

单价法是指以项目的单个单元为计价的预算方法，适合于过程比较固定、重复性较大的项目类型。最典型的使用单价法进行预算的项目是地图数字化，例如要对 500 幅地形图进行数字化，在考察大多数地形图内容、质量、预期数据量后，根据最终确定的数字化流程以及精度要求，可以推算出平均每幅地形图所需要的时间和费用，进而就可以得到总的项目费用。

6．2．4 根据项目参与人员的费用做预算的方法

根据项目参与人员的费用来做预算取决于项目使用的人员类型，通常该方法在使用高级技术人员的情况下使用。该方法注重每个人参加项目的时间，而无需对项目的任务加以严格定义。

如同进度安排一样，进行项目预算时，为了能够应付项目过程中出现的各种问题，保持一定的“弹性”是必需的。

7．人员管理

7．1 地理信息系统应用项目中的人员

在地理信息系统应用项目的组织和建设过程中，不论是由应用部门独立开发还是和由专门的独立软件开发商进行开发，其中涉及的人员按照角色，可以分为以下几类：

7．1．1 项目管理人员

项目管理人员负责整个项目，又称为项目经理。在项目实施之前，项目管理人员参与制

订工作计划；在项目开发过程中，要监督和管理项目的进度与开销。当进度落后或者开销超出预算时，项目管理人员要找出其中的原因，并提出解决的方案。

在具体实践中，根据项目的规模以及开发组织方式（合作开发还是单独开发），可以设一个或多个项目管理人员，具体分管不同的方面。项目管理人员应该对项目目的以及相关技术相当熟悉，并且具有一定的管理能力。在合作进行的项目开发中，一般由双方各出一人担任项目管理人员，以进行协调工作。

7．1．2 系统开发人员

系统开发人员在应用项目的开发过程中，所占比例应该是最大的。按照所进行开发活动，可以分为系统分析员、设计员、编程实现人员、测试人员以及系统维护人员。同时可以按照项目的具体技术内容，划分为几个技术专题，每个专题对应一个或多个开发人员。

一般而言，系统分析员、设计员、编程实现人员不应该重叠，但是在实际过程中，由于人力资源不足，更由于没有很好地实施软件工程，分析设计文档不够完善，造成交流困难，一般都按照技术专题对人员进行分组，每一组完成该专题的需求分析，设计和实现工作。在项目规模较小时，可以采用这种人员组织方式，以提高灵活性。

7．1．3 数据录入和处理人员

数据在 GIS 应用项目中占有非常重要的位置。一个 GIS 项目，往往需要大量的数据，所以在进行软件开发的同时，需要组织数据的录入和处理。数据录入和处理没有特殊的技术要求，一般的工作人员，经过简单的培训即可胜任该工作。在项目进行中，需要考虑的因素有：

1）数据录入和处理是一件相当单调而且令人厌烦的工作，需要工作人员认真负责，同时还要采取适当的激励手段以保证工作顺利进行；

2）在数据录入和处理过程中，需要专门的技术人员进行质量检查，避免数据精度不能满足最终项目的要求；

3）如果待开发软件包括了数据录入功能，如果等待软件完成后再进行数据录入，可能会造成工期延误，因此可以考虑先采用其它软件进行数据录入和处理，然后再进行数据转换。

7．1．4 开发支持人员

开发支持人员保证整个开发过程的顺利进行，其工作包括开发工具的维护，网络管理等等，如果项目规模较大，需要配置管理人员保证协作开发不至于引起混乱。

7．1．5 领域专家

由于许多 GIS 应用项目对应于专业的要求，所以需要领域专家参与开发过程，以便于将他们的理论、经验在计算机系统中加以实现。

7．1．6 用户

用户不直接参与开发过程，但是他们是系统的最终使用者。在开发系统之前，需要从用户那里获得需求；在开发过程中，完成的原型系统一般需要由用户进行评估。

7．1．7 其它支持人员

包括不直接参加项目的支持人员，如财务人员，文档管理人员等等，他们对于项目的正常运行起着重要的作用。

7．2 对人员需求的了解——进行管理的基础

在项目过程中，需要对参加项目的各类人员进行妥善的安排，对于人的管理，首先要求项目管理人员了解各个执行人员的性格特性，从而了解其需求*，找到调动其积极性的因素，并积极满足其需求。在信息技术领域，员工的积极性直接影响了工作效率和质量，因而，在GIS项目组织管理中，调动员工积极性对于项目管理人员尤为重要。

一般认为，与工作有关的需求因素包括：

1）工资或奖金的多少；

2）在机构内的职位和责任范围；

3）人际关系；

4）工作的挑战性；

5）个人的被承认程度；

6）所拥有的自由度；

7）享受各种待遇的程度，如医疗保险、假期等等；

8）物质方面，包括办公条件等等。

不同类型的人对不同需求的要求程度也有所差异，如工作型的人员期望得到较好的工资、待遇、职称、办公条件，而事业型的人员则更看重工作的挑战性、责任范围、个人的发展等方面的内容。此外，人员的性格也决定了相互之间的合作模式以及他们在组织中最合适的位置。不同性格的人可以合理地安排在一起，形成一种和谐的工作环境，保证项目的顺利完成。

在工作中，通常高的收入是工作人员最基本的需求，换言之，提高工资报酬是管理人员采取的主要激励手段。但是，经济学家认为，劳动力的供给是“向后拐弯”的曲线（图 15-8）。

---

* 马斯洛的需要五层次包括：生理的需要，安全的需要，爱和归属的需要，获得尊重的需要，自我实现的需要。人员在工作中的需求可以归纳到这五个层次中。

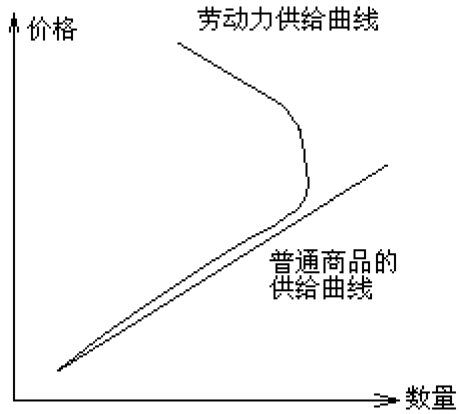


图 15-8：劳动力供给曲线

造成种现象的主要原因是，当雇员的工资达到一定水平之后，随着收入的增长，追求更多的闲暇，使工作时间减少。对于项目管理人员，了解劳动力供给的这种特性，可以采取更加灵活的刺激手段，例如对于开发人员，一个具有挑战性的工作往往能够激发其工作热情。

7．3 人员的组织管理

项目管理者除了根据每个工作人员的具体情况确定其分工，采用以下管理手段有助于提高组织的生产效率：

1）工作的内容

保证分配适当的任务给雇员，其分配依据包括其技术与性格的特点。此外，一般要给出较多的工作量，使得每个工作人员总是有足够的工作做。

2）工作的安排

根据项目所需要的时间，每隔一定时间对整个任务进行安排。

3）交流

与技术人员的交流，可以使管理人员了解项目的具体进度及遇到的技术难点，还可以使技术人员把握整个项目的技术、时间、质量等具体的要求。

4）项目背景和发展要求的回顾

给技术人员较为全面的项目介绍，包括背景、参与的团体、进度安排等，容易使技术人员了解其工作在整个项目中具体的位置，增强其成就感。

5）层次

在关键的岗位使用优秀的人才。

6）新技术

关注新技术的发展，注重技术革新，通过组织技术培训保证技术人员掌握新技术。

7）会议

适当地组织不同层次、规模的会议，在会议进行之前，组织者要明确会议的主题和目的，使参与人员有所准备。

8）交流手段

采用电子邮件是较好的信息交流手段，它可以作为开发的历史档案加以保存。

在信息技术的开发和应用中，人员的管理是企业管理中最重要的一个环节，对于 GIS 应用项目，由于涉及各个领域的技术人员，更需要有效的人员组织和管理。

8．开发和数据管理

在建立地理信息系统应用的过程中，软件开发和数据处理是最为繁重的两个技术任务，对这两个部分的管理的好坏直接影响到整个项目的成败。

8．1 软件开发管理

尽管在软件工程方面已经有许多的技术和方法，如结构化方法，面向对象的方法等等，在开发过程中采用这些方法以提高效率，但是开发过程的管理始终是决定一个软件项目成败的重要因素。对于开发管理，并没有“万能”的办法可以适用于所有的软件开发过程，但是在管理中遵循以下原则可以更好地确保软件产品质量和开发效率[B.W.Boehm]，这些原则同样也适用于 GIS 平台软件的开发。

1）用分阶段的生命周期计划严格管理

在整个软件生命周期中，要完成许多性质各异的工作。该原则意味着把软件生命周期分成若干个阶段，并制定严格计划对软件开发和维护进行管理。通常以下六类计划是应该制定并严格执行的，即项目概要计划、里程碑计划、项目控制计划、产品控制计划、验证计划和运行维护计划。

2）坚持进行阶段评审

软件的质量保证不能等到编码结束之后再进行，这是因为，第一，大部分错误是在编码之前造成的；第二，错误发现越晚，所需要代价越高。因此，需要在每个阶段都进行严格评审，以便尽早发现软件开发过程中的错误。

3）实行严格的产品控制

在开发过程中，用户的需求难免发生变化，而为了满足其变化的需求往往要付出较高的代价，这就需要进行严格的产品控制，通常所采用的技术就是配置管理。

4）采用现代程序设计技术

实践表明，采用先进的程序设计技术既可以提高开发效率，又可以提高软件的可维护性。这些技术包括结构化分析、设计、编程技术以及面向对象的分析、设计、编程技术等。

5）结果应能够清楚地审查

软件产品是思维的产物，这决定了工作过程难于评价和管理，难以准确度量。为了提高开发过程的可见性，以便进行管理，应该根据总目标和进度，规定开发组织的责任和产品标准。

6）开发小组的人员应该少而精

软件开发小组的组成人员素质应该好，而人数不宜过多。素质高的人员开发效率比素质低的人员的开发效率可能高几倍到几十倍，而且错误更少。而小组人数的增多会造成交流的困难。

7）承认不断改进软件工程实践的必要性

由于技术在不断进步，应该主动采纳新技术，而且要不断总结经验，以不断改进工程实践*。针对一个项目，遵循前六点就可以保证项目的成功，该原则使软件组织能够持续发展。

---

* 这也正是CMM第五级所强调的。

8．2 数据管理

在一个 GIS 应用项目中，通常需要输入和处理大量的数据，其中主要是空间数据，如果没有采取合适的数据管理，那么可能因为以下原因造成项目的混乱甚至失败：

1）数据质量不能达到项目要求；

2）数据没有完全输入或处理，造成项目延期；

3）数据重复录入或处理，造成人员和时间的浪费；

4）进行数据处理和模型运算时，没有及时采用最新的数据。

为了避免出现上述的问题，必须对数据的录入和处理过程和数据质量进行严格的控制。通常 GIS 应用项目的空间数据来源包括：GPS 数据，摄影测量数据，测量数据，卫星遥感数据，已有非数字地图。不同的来源，其处理和质量控制原则也有所差别。对于 GPS 数据和测量数据，只需要进行简单的数据规范化（如格式转换，拓扑关系建立等等）处理，就可以在 GIS 中使用，其精度由 GPS 或测量设备决定，在数据处理过程中不会带来新的误差。摄影测量数据和卫星遥感数据通常需要进行一系列纠正操作，然后进行自动或半自动解译，以提取专题信息，最后通常是转换为多边形数据，输出到 GIS 应用中，其误差来源包括纠正（包括辐射纠正和结合纠正）以及分类过程中的误分。纸质地图的录入，可以采取手扶跟踪数字化和扫描矢量化两种方式，其误差来自于数字化过程，并且与进行数字化人员的工作态度有密切的关系，此外数字化数据往往现势性较差。

数据输入和处理的进度控制可以采用普通的里程碑方法或甘特图法，如果工作区范围较大，并且每一数据处理单元只覆盖工作区的一部分，可以采用图 15-9 所示的方法，以表示进度，其优点是更加直观，不容易遗漏数据。

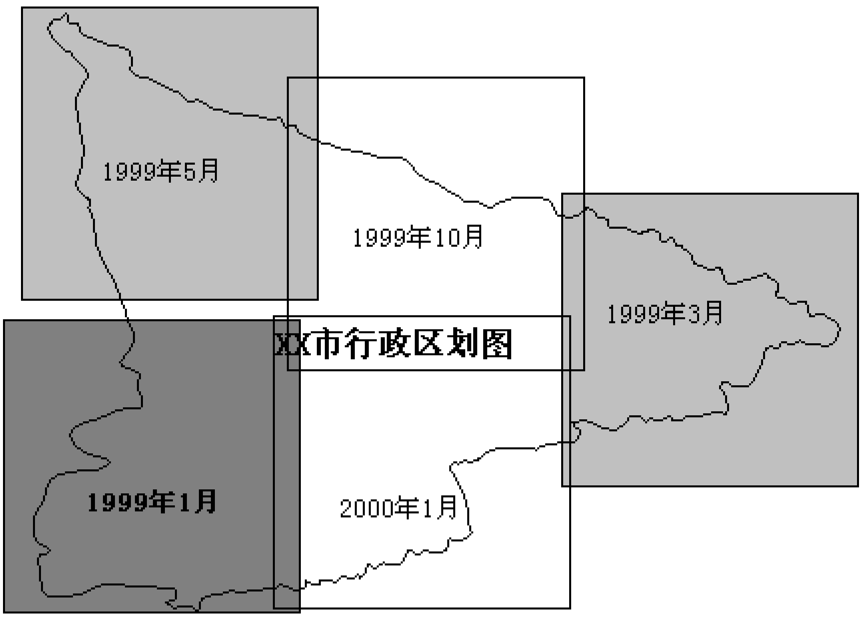


图 15-9：一种直观的数据输入进度管理方法示例

（图中每一矩形框内的时间表示该数据应该完成的时间，

深灰色矩形表示已经完成，而浅灰色矩形表示数据录入和处理正在进行）

如果项目中数据量大，并且需要多个数据文件以存储数据，那么需要在工作进行之初就定义文件的命名规则，以通过文件名表达其数据内容，避免造成管理上的混乱*。如用文件

* 采用地理信息系统软件工程技术一章中描述的数据配置管理方法可以更好的避免数据管理混乱。

名SO01DG表示第一次数字化的土壤图（SO：Soil，土壤；01：第一次；DG：数字化）。通过文件名表达的信息可以有专题、已进行的处理、时间、处理人员、数据区域范围等等，但是太长的文件名同样会造成记忆的困难。

除了定义文件命名规则外，其它描述信息，包括数据来源、地物的数值范围、地图投影、精度、数据的依赖关系等元数据可以建立数据字典。可以对数据字典进行检索，以掌握数据的进度。

9．项目控制与评估

在 GIS 应用项目进行中，控制活动要贯穿于整个项目的始终，才不至于在项目完成后出现严重的质量以及时间问题。而在项目的每一个阶段、或者整个完成之后，进行事后分析和评估可以更好地了解项目进展中的问题，对项目参与人员进行评定，为下一阶段或者以后类似项目组织管理作为参考依据。从一定意义上来讲，进行评估是实施项目控制的有效途径。

通常，进行事后分析和评估的对象包括：

1）程序：数量和质量；

2）数据：数量和质量；

3）参与人员：工作时间和效率；

4）资金：使用情况和使用效率；

5）设备：占用情况和效率。

在对上述因素进行分析和评估时，要尽可能使用定量的方法，一方面数量能更准确地说明项目进展和存在的问题；另一方面，采用数量也便于进行横向和纵向的对比。横向对比是指项目中各个小组，甚至不同人员的对比，纵向对比是指项目的不同阶段或不同项目之间的对比。

分析评估的结果可以以表格、统计图的形式直观的表现（表 15-6，图 15-10），项目管理人员可以根据结果评估工作进展，发现问题，并在下一步的工作中进行改正。

表 15-6：XX 项目第一阶段（1999 年 1 月-1999 年 10 月）工作分析结果

项目	数值
程序	
代码总行数	78532
错误数目	130
错误比率	1.65 处/千行
数据	
录入总数据量	78 兆
抽样平均偏移误差	0.12 毫米
遥感分类后数据总量	290 兆
抽样分类精度	83%
人员	
人员总工作时间	140 人月
资金	
总使用资金	8.2 万元

表 15-6 列出了从总量上对项目中各个要素分析的结果，可以作为计算项目阶段成本、效率的依据；而图 15-10 更好的说明了项目的进展状况，可以使管理者方便地发现项目进展顺利和不顺利的时间段，并进而发现其原因，作为下一阶段工作管理的依据。而整个项目的分析图表可以被复用于下一个类似的项目，进行成本——效益分析和进度安排。

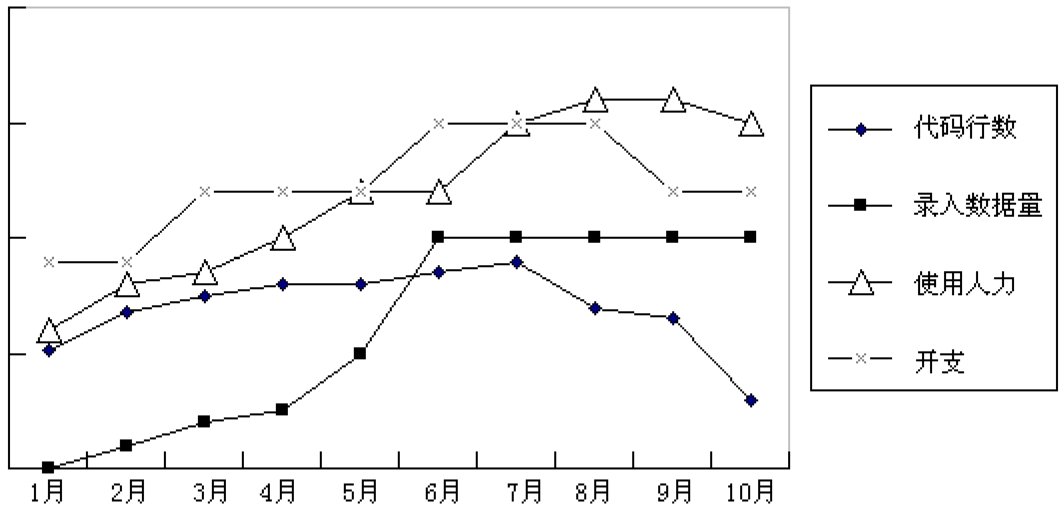


图 15-10：通过不同指标反映项目的进展状况

10．软件研制和开发的质量管理——对于 GIS 软件开发组织的指导

由于计算机技术的发展以及用户需求的增长，GIS 软件的研制与开发（R&D，Research and Development）也越来越多，这意味着竞争也愈加激烈。对于 GIS 软件开发组织，在市场上取得成功的因素包括：技术的先进性、功能的实用性和可靠性、系统操作的便利性、版本的连续性以及技术支持维护等等。为了达到这些目的，唯一的途径就是采用合适的质量管理与控制。

10．1 ISO-9000 系列标准

质量管理的目的是为了生产高质量的产品，它不是通过检验产品的质量，而是把质量建立在生产过程之中来达到。近年来，国际上影响最为深远的质量管理标准是 ISO9000 系列标准。ISO9000 最初主要针对制造行业，现在已经扩展到硬件、软件甚至服务领域。ISO9000 强调控制、可审查性、检验与核实以及过程的改进。其基本要求是："说你做的(Say what you do)"，"做你说的(Do what you say)"，并"证明你已做的(Demostrate what you have done)"。

ISO-9000 系列标准包括：

1）ISO-9000 质量管理和质量保证标准——选择和使用的导则；

2）ISO-9001 质量体系——设计/开发，生产，安装和服务中的质量保证模式；

3）ISO-9002 质量体系——生产和安装中的质量保证模式；

4）ISO-9003 质量体系——最终检验和测试中的质量保证模式；

5）ISO-9004 质量管理和质量体系要素——导则。

ISO-9000 原本为了制造业而规定，不能直接应用于软件行业，后来追加了 9000-3 标准，成为"使 ISO-9001 适用于软件开发、供应和维护的指南"。

ISO-9000-3 的核心思想是"将质量制作入产品之中"，也就是说，软件的质量提高不能依赖于完成后的测试，而取决于整个软件生存期所有的活动。ISO-9000-3 的要点如下：

1）ISO-9000-3 标准不适用于面向多数用户销售的程序包软件，仅适合于依照合同进行

的单独订货开发软件，它也是用户企业的系统部门在建立质量保证体系时的指南；

2）ISO-9000-3 标准对供需双方领导的责任都做了明确的规定，并没有单纯地把义务全部加在供方；

3）在包括合同在内的全部工序中进行审查，并彻底文档化。具体来说，就是需方和供方合作审查，找出含混不清的问题，并记录在文件中；

4）ISO-9000-3 叙述了供方和需方如何合作进行有组织的质量保证活动才能制作完美的软件，强调质量保证体系贯穿于整个开发过程，强调防患于未然而不是事后纠正；

6）供方应实施内部质量审核制度，要求供方为了质量管理而整顿其组织机构，建立监督质量体系的机制；

7）供方应对每项合同进行审查。

10．2 软件过程改进的 CMM 模型

由于 ISO-9000 系列标准并不针对于软件企业，到九十年代初，CMU 的 SEI(软件工程研究所)提出了软件的能力成熟度模型(CMM-Capability Maturity Model for Software)。

与 ISO9000 相比，CMM 只关注软件，明确强调持续的过程改进，而 ISO9000 只是解决质量体系的最小保证。

CMM模型较为全面地描述和分析软件过程*能力的发展程度，建立了描述一个组织的软件过程成熟程度的分级标准。利用该标准，软件组织可以评估自己当前的过程成熟程度，并通过提出更严格的软件质量标准和过程改进，来选择自己的改进策略，以达到更高一级的成熟程度。CMM模型共有五个成熟度级别，其特征为：

1）初始级(Initial)

软件过程杂乱无章，有时甚至混乱，几乎没有明确定义的步骤，成功完全依赖于个人努力和英雄式的核心人物。

2）可重复级(Repeatable)

建立了基本的项目管理过程来跟踪成本、进度和机能，有必要的过程准则来重复以前在同类项目的成功。

3）定义级(Defined)

管理和工程的软件过程已经文档化、标准化，并综合成为整个软件开发组织的标准软件过程，所有的项目都采用根据实际情况修改后得到的标准软件过程来开发和维护软件。

4）管理级(Managed)

制订了软件过程和产品质量的详细度量标准。软件过程和产品的质量都被开发组织的成员所理解和掌握。

5）持续优化级(Optimizing)

加强了定量分析，通过来自过程质量反馈和来自新观念、新技术的反馈使过程能够不断持续地改进。

CMM 定义了每个成熟度级别的关键过程范围（图 15-11），一个软件组织只有满足该范围内的每个目标，才能被认为达到该成熟度级，这实际上也是软件组织达到更高一级成熟度的指导。

---

* 软件过程是人们用于开发和维护软件及其相关产品的一系列活动、方法、实践和改造。软件的开发有两种截然不同的方式：“面向过程的（Process Oriented）”和“面向责任的（Commitment Oriented）”，前者强调过程的规范和持续优化，后者更强调个人的能力。目前许多软件企业都开始注重软件过程。

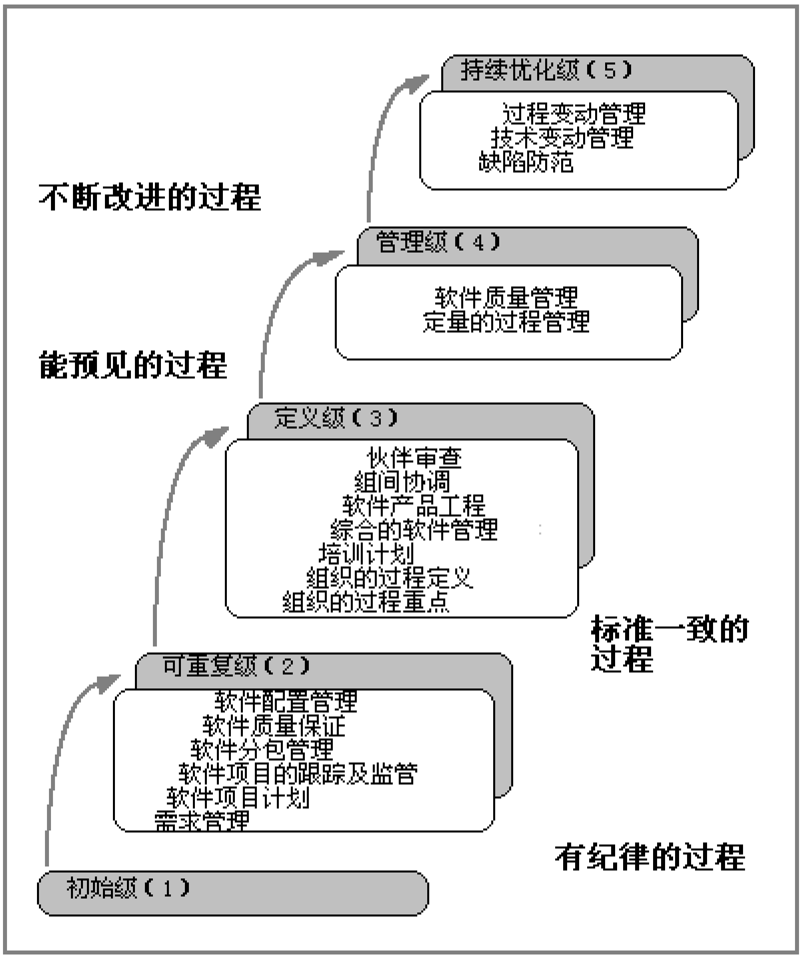


图 15-11：CMM 模型中的五个成熟度级别，其关键过程范围[CMU SEI]

CMM 建立了一个可用的标准描述，在签定软件项目合同时，可以参考这些标准进行风险评估，而软件企业可以利用 CMM 框架来改进其开发和维护过程。

尽管 CMM 列出了每个成熟度级别的特征和关键过程，但是并没有回答如何做才能达到高一级的成熟度级的问题，没有提出特定的软件技术，也没有涉及如何具体的进行人员管理，这些都需要应用 CMM 的组织在实践过程中，根据其具体情况加以解决。

工欲善其事，必先利其器。

孔子

第十六章 地理信息系统软件工程技术

导读：本章介绍了软件工程的基本概念，软件开发活动，过程模型以及两种重要的开发方法：结构化方法和面向对象的方法。在 GIS 开发过程中应用软件工程技术，可以提高软件开发效率和质量。本章讲述了软件工程技术在 GIS 系统开发中的几个应用方面:

GIS 软件构件和构架

需求分析

数据管理设计

用户界面设计

设计模式在 GIS 软件开发中的应用

应用配置管理来实现空间过程支持

这些方面涉及了 GIS 开发过程中不同的阶段以及不同的层次，有些方法之间是互斥的，如 UML 和 Code 方法，但是软件工程技术最重要的是实用，开发者可以根据具体的情况选用不同的技术。

1. 软件工程简介

1．1 基本概念

计算机软件工程是一类求解的工程。它应用计算机科学、数学及管理科学等原理，借鉴传统工程的原则、方法，创建软件以达到提高质量，降低成本的目的。其中，计算机科学、数学应用于构造模型与算法，工程科学用于制定规范、设计范型，评估成本及确定权衡，管理科学用于计划、资源、质量、成本等管理。从学科角度来看，软件工程是一门指导计算机软件开发和维护的工程学科。

软件工程的提出，是为了解决 60 年代出现的软件危机，当时在大型软件开发中存在着价格高，开发不容易控制，软件开发工作量估计困难，软件质量低，项目失败率高等许多问题，给软件行业带来了巨大的冲击。软件工程的研究，提出了一系列理论、原则、方法以及工具，试图解决软件危机。

和其它工程一样，软件工程有其目标、活动和原则，其框架可以概括为图 16-1 所表示的内容。

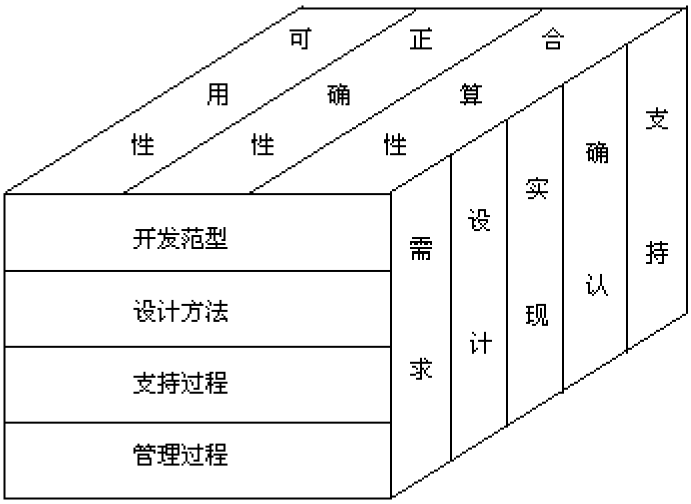


图 16-1：软件工程框架[王立福]

软件工程的目标可以概括为“生产具有正确性、可用性以及开销合宜的产品”，其活动包括需求、设计、实现、确认以及支持等活动，围绕工程设计、支持以及管理，有以下的四条基本原则：

1）选取适宜的开发模型，可以认识需求易变性，并加以控制，以保证软件产品满足用户的需求；

2）采用合适的设计方法，通常要考虑实现软件的模块化、抽象与信息隐蔽、局部化、一致性以及适应性等特征；

3）提供高质量的工程支持，在软件工程中，软件工具与环境对软件过程的支持颇为重要；

4）重视开发过程的管理，软件工程的管理，直接影响可用资源的有效利用、生产满足目标的软件产品，提高软件组织的生产能力等问题。只有当软件过程予以有效管理时，才能实现有效的软件工程。

1．2 软件工程活动

软件工程活动包括需求、设计、实现、确认及支持等，它们对应于软件开发活动的不同阶段，一般来说，软件开发都要经历从分析设计到实现确认的过程。在每个阶段按照相应的规范进行工作，并得到该阶段的成果，是保证整个开发活动成功的关键。

1．2．1 需求分析

需求分析阶段处于软件开发的前期，其基本活动是准确定义未来系统的目标，确定为了满足用户的需求必须做什么。需求分析又划分为两个阶段，即需求获取和需求规约，前者是用自然语言清楚地描述用户的要求，而需求规约的目的是消除获取需求的二义性和不一致性。

在软件生命周期中，一个错误发现得越晚，修复错误的费用也越高，所以，高质量的需求工程是软件项目得以正确、高效完成的前提。对于系统分析人员，建立需求面临着以下三个方面的困难：

1）问题空间的理解，系统开发人员通常是计算机专业人员，难以深入理解各种业务系

统所要解决的问题空间；

2）人与人之间的通信，对于系统分析人员而言，通信主要包括同用户的通信以及同事之间的通信，由于自然语言的二义性，会给准确刻画需求造成障碍；

3）需求的不断变化，造成需求变化的原因很多，包括技术，用户方，市场等等，作为分析人员，必须采用一些策略以适应变化。

面向对象的分析方法被认为是解决上述困难较好的技术，但是完整、准确的刻划问题空间始终是分析人员所面临的挑战。

1．2．2 系统设计

一般来说，需求分析阶段的主要任务是确定系统“做什么”，而设计阶段则要解决“怎么做”的问题。通常*设计阶段又划分为总体设计和详细设计，总体设计确定系统的总体结构框架；而详细设计要具体地描述如何具体地实现系统，通常可以依据详细设计的结果进行编码。详细设计包括：详细的算法；数据表示和数据结构；实施的功能和使用数据之间的关系。详细设计过程中，采用了一些工具，以便对数据、算法等进行描述，包括流程图，PAD(Problem Analysis Diagram)，盒图(N-S图)，伪码等等。

1．2．3 实现阶段

在软件实现阶段，要将设计的结果变换成程序设计语言编写的程序。在实现阶段，首先要确定程序设计语言，其影响因素包括：开发人员对语言的熟悉程度，语言的可移植性，编译程序的效率，编译工具的支持等等。目前，C++语言是普遍被采用的构造系统软件的编程语言，而 Java 则更多地应用于编写网络程序。

无论采用哪一种编程语言，都要求编写高质量的源程序代码，程序质量通常包含正确性、可读性、可移植性、程序效率等指标。考虑到系统的维护和演化，提高源程序的可读性是实现阶段的一个重要目标，其途径包括添加注释，规范书写格式，确定标识符命名原则*，采用结构化的程序设计（不用或减少使用goto语句）等等。

1．2．4 确认活动

尽管确认活动贯穿于软件开发活动的始终，但是系统完成后的软件测试是主要的确认活动。软件测试是指按照特定规程，发现软件错误的过程。软件测试的技术大体上可以分为两类，即白盒测试技术和黑盒测试技术，前者依据的是程序逻辑结构，后者依据的是软件行为描述。根据测试的步骤，测试活动又可以划分为单元测试，集成测试，确认测试和系统测试，其中确认测试是为了检验软件的功能和性能是否与用户需求一致，而系统测试主要是测试软件同硬件、其它支持软件、数据等结合在一起，在实际的运行情况下，同用户需求的匹配程度。

---

* 这里实际上是指结构化的设计，在面向对象方法学中，各个阶段的分界已经被模糊化了。
* 微软所采用的匈牙利命名法是一个较好的对标识符进行命名的方法。

1．2．5 软件维护

当软件开发完成并交付用户使用后，就进入运行/维护阶段，在运行/维护阶段仍需要对软件进行修改，称为软件维护，软件维护活动可以分为以下几类：

1）改正性维护，其目的是为了纠正运行阶段发现的软件错误，性能上的缺陷以及排除实施中的误用。

2）适应性维护，随着计算机的发展，软件的外部环境或者数据环境发生变化，为了使之适应这种变化而对软件的修改称为适应性维护。

3）完善性维护，在使用过程中，用户往往会对软件提出新的功能和性能需求，为了满足这些需求，需要修改或再开发软件，称为完善性维护。

4）预防性维护，预防性维护的目的是为了提高软件的可维护性，可靠性等，为进一步的软件维护打下良好的基础。预防性维护一般由开发单位主动进行。

1．3 结构化方法和面向对象方法

在进行系统分析设计的过程中，逐渐形成了一些系统化的方法，以便于更好地描述问题域及进行系统设计，目前经常采用的两种方法是结构化方法和面向对象的方法。

1．3．1 结构化分析和设计

结构化的方法基于模块化的思想，采用“自顶向下，逐步求精”的技术对系统进行划分，分解和抽象是它的两个基本手段。

结构化分析将软件视为一个数据变换装置，接受各种输入，通过变换产生输出。数据流图（DFD，Data-Flow Diagram）是一种描述数据变换的工具，是结构化分析普遍采用的表示手段。数据流图由五个部分组成，即：加工，数据流，数据存储，数据源和数据潭，其中数据流表示数据和数据的流向，而加工是对数据进行处理的单元（图 16-2）。除了数据流图以外，还需要数据字典和说明分别对数据流和加工进行描述。

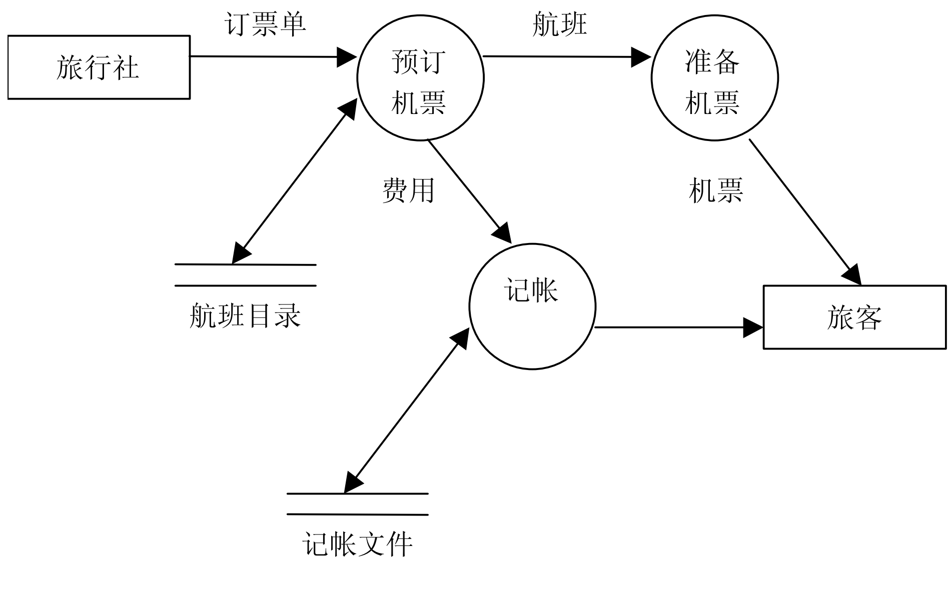


图 16-2：一个典型的数据流图：飞机订票

结构化的系统总体设计主要是确定模块结构图，以描述功能模块之间的关系，一些主要的表示形式有：层次图，HIPO（层次+输入/处理/输出）图，结构图等等。结构化的设计定义了一些原则和方法，可以将数据流图“映射”成为模块结构图。

1．3．2 面向对象的分析和设计

面向对象的方法学认为，客观世界是由许多各种各样的类组成的，每种对象都有各自的内部状态和运动规律，对象之间的作用和联系就构成了各种不同的系统。面向对象方法学所追求的是使解决问题的方法空间与客观世界的问题空间结构达成一致。由于面向对象的技术在理解问题空间、控制需求变化、消除从分析设计到编码的“鸿沟”、支持软件复用等各个方面优于其它方法，使之称为目前软件开发的主流方法。

目前已经提出了多种不同的面向对象的分析、设计方法，如 Cord-Yourdon 方法，Booch 方法，OMT 方法，Jacobson 的 use case 驱动方法等等，这些方法在侧重点、符号表示和实施策略上有所不同，但是其基本的概念是一致的，这些概念有：对象，类，属性，服务，消息，继承，封装等等。

近年来，综合Booch方法、OMT方法以及use case的UML(Unified modeling language，统一建模语言)逐渐成为主要的面向对象方法。图 16-3 给出了用UML表达几何体的例子*。

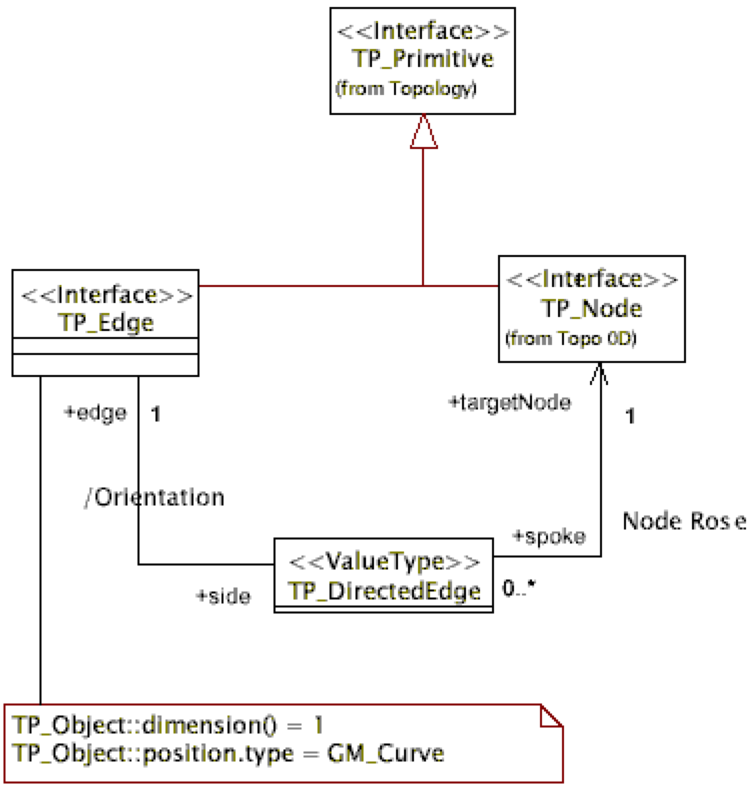


图 16-3：用 UML 表达的线几何体类以及和其它类的关系（OpenGIS Consortium）。

* 图的具体含义请参阅ULM和OpenGIS的相关规范。

1．4 开发过程模型

软件开发模型是软件开发全部过程、活动和任务的结构框架。软件开发模型能够清晰、直观的表达软件开发过程，明确规定要完成的主要活动和任务，可以作为软件项目工作的基础。

随着软件工程了实践，相继提出了一系列开发模型，如下：

1．4．1 瀑布模型

在瀑布模型中，将各项活动规定为依照固定顺序连接的若干阶段工作，形如瀑布流水（图16-4），瀑布模型的特征是：每一阶段接受上一阶段的工作结果作为输入；其工作输出传入下一阶段；每一阶段工作都要进行评审，得到确认后，才能继续下阶段工作。瀑布模型较好地支持结构化软件开发，但是缺乏灵活性，无法通过软件开发活动澄清本来不够确切的需求。

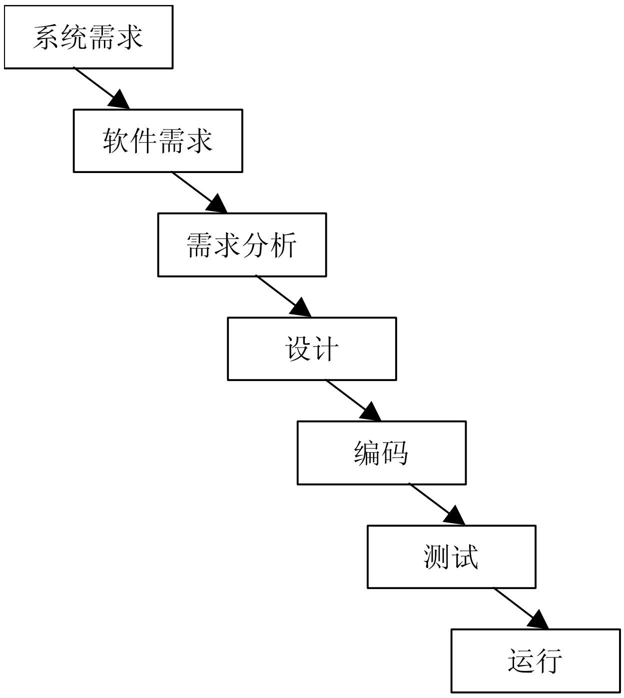


图 16-4：瀑布模型

1．4．2 演化模型

演化模型主要针对事先不能完整定义需求的软件开发。用户可以先给出核心需求，当开发人员将核心需求实现后，用户提出反馈意见，以支持系统的最终设计和实现。

1．4．3 螺旋模型

螺旋模型是在瀑布模型以及演化模型的基础上，加入风险分析所建立的模型。在螺旋模型每一次演化的过程中，都经历以下四个方面的活动：

1）制定计划——确定软件目标，选定实施方案，弄清项目开发的限制条件。

2）风险分析——分析所选方案，考虑如何识别和消除风险。

3）实施工程——实施软件开发。

4）客户评估——评价开发工作，提出修正建议。

每一次演化都开发出更为完善的一个新的软件版本，形成了螺旋模型的一圈。螺旋模型借助于原型，获取用户需求，进行软件开发的风险分析，对于大型软件的开发，是颇为实际的方法。

1．4．4 喷泉模型

喷泉模型体现了软件开发过程中所固有的迭代和无间隙的特征（图 16-5）。喷泉模型表明了软件刻画活动需要多次重复。例如，在编码之前，再次进行分析和设计，并添加有关功能，使系统得以演化。同时，该模型还表明活动之间没有明显的间隙，例如在分析和设计之间没有明确的界限。

在面向对象技术中，由于对象概念的引入，使分析、设计、实现之间的表达连贯而一致，所以，喷泉模型主要用于支持面向对象开发过程。

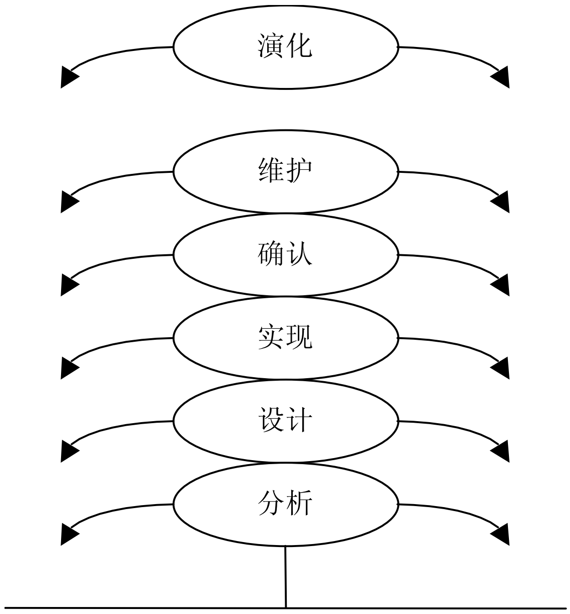


图 16-5：喷泉模型

目前，随着面向对象技术的发展和 UML 建模语言的成熟，统一软件开发过程(USDP，Unified Software Development Process)被提出以指导软件开发，它是一个用例（use case）驱动的、体系结构为中心的、增量迭代的开发过程模型，适用于利用面向对象技术进行软件开发。

2．GIS 领域的体系结构和构件

按照应用目的，地理信息系统可以分为区域地理信息系统、专题地理信息系统以及地理信息系统工具，它们共同组成了 GIS 领域（Domain）。所谓领域，是指共享某种功能性(Functionality)的系统或应用程序的集合，换言之，领域表现了一组应用系统共性的方面。对于地理信息系统而言，其共享的功能就是对空间数据输入、管理、分析和表现，而这恰好是地理信息系统工具所提供的功能（图 16-6）。



图 16-6：GIS 领域

与具体应用领域相对应，对于领域，可以实施领域工程，得到领域模型，并建立领域特定的软件体系结构（DSSA, Domain-Specific Software Architecture）。DSSA 是能够适应领域中多个系统的需求的一个高层次的设计，在开发具体应用系统时，可以利用 DSSA 将领域构件（Component）连接和组织起来，以支持软件复用，提高软件开发的效率和质量。

对于地理信息系统，实施领域工程，得到 DSSA，不仅可以指导具体应用系统的开发，而且 DSSA 可以直接作为地理信息系统工具的软件体系结构。

地理信息系统的核心功能包括空间数据的输入、管理、分析以及表现，并且这些功能形成了一个比较完全的数据处理流程，此外考虑到与遥感以及全球定位系统的结合，形成如下的系统结构（图 16-7）。该视图更多地体现了地理信息系统的业务逻辑，为了适用于具体的应用系统，该结构可以被特化——类似于面向对象中，从父类派生一个子类——形成更加具体的体系结构。

依赖于 DSSA，可以将构件组装起来，形成具体的应用系统，基于构件的技术已经成为软件开发技术的主流，它从面向对象技术发展而来，是开发高效、低成本程序的重要实现途径。

为了能够通过组装以构造系统，构件必须能够互相合作，即具有互操作性，这是通过定义构件的接口规范来实现的。对于构件而言，除了互操作性之外，还要支持分布式的网络计算，即构件的互操作可以是基于异种平台的，其实现需要分布计算平台（DCP-Distributed Computing Paltform）的支持。

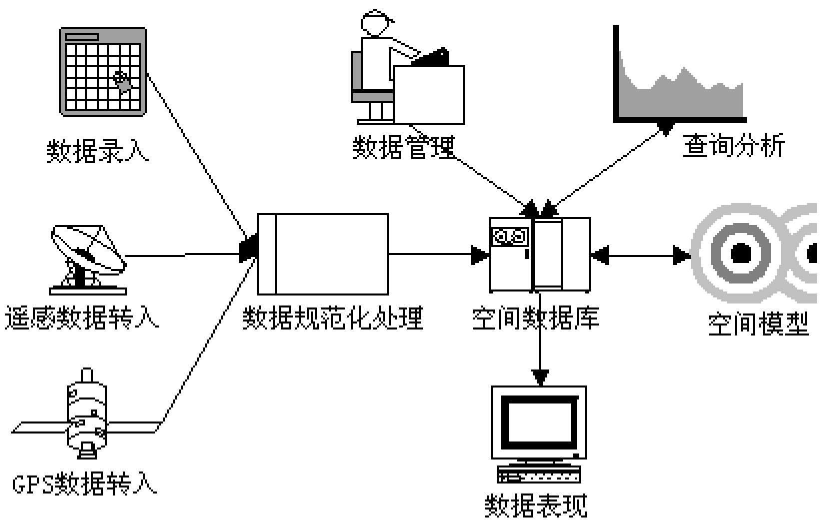


图 16-7：GIS 领域体系结构：一个工作流视图

目前存在着多种构件技术标准，其中 OMG（对象管理组织）的 CORBA（公共请求对象代理体系结构-Common object request broker architecture）和 Microsoft 的 OLE/COM/DCOM 技术是其中两个主要的、被广泛采用的标准。

COBRA 定义了一个带有开放软总线的分布式结构，在这一结构中，来自不同厂商、运

行于不同操作系统上的对象，能够进行互操作。CORBA 对象的互相通信通过对象请求代理（ORB，Object Request Broker）为中介，可以在多种流行网络通信协议上实现。接口描述语言（IDL，Interface Description Language）用于描述对象接口，它与语言无关，使得所有 CORBA 对象以一致的方式被描述。

Microsoft 的 DCOM（分布式对象构件模型，Distributed Component Object Model）技术是对原有的 COM 技术的扩展，以支持在网络上不同计算机的对象之间的通信。COM 定义了接口的二进制标准，包括接口交互、管理对象及其资源等等。而 DCOM 通过增加网络协议的支持，使得对象可以通过网络互操作。DCOM 技术很好的支持复用，位置独立，可扩展等，并且其执行性能较好，目前被基于 Windows 平台的软件开发商所广泛支持。

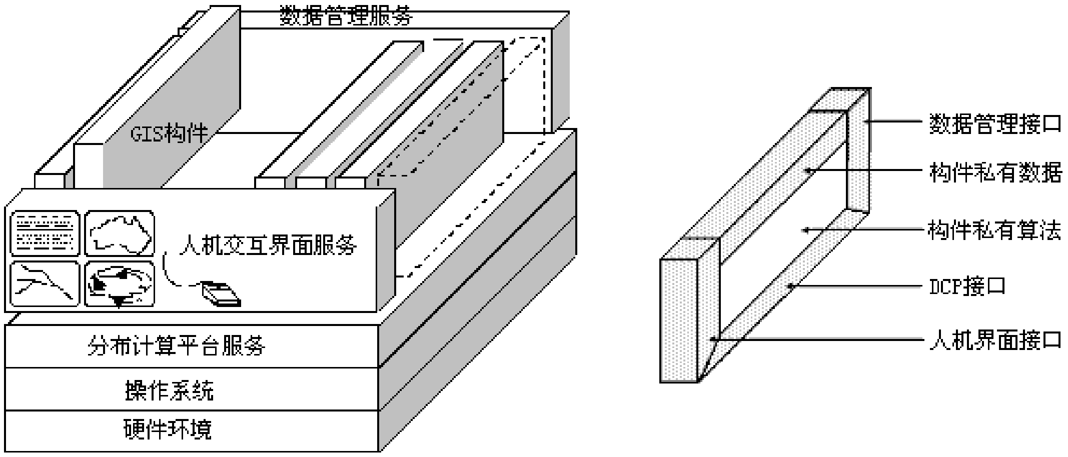


图 16-8：基于 DCP 的 GIS 体系结构和构件：一个实现视图[OpenGIS Consortium]

对于GIS软件开发，构件技术带来的除了实现复用、支持分布的互操作外，还具有以下两个方面的优势*：

1）数据是 GIS 中最为重要的部分，目前各个软件数据格式不同，不便于共享和保护用户已有投资，构件技术通过互操作的支持可以实现数据共享；

2）地理信息系统功能的涵盖范围非常广泛，包括图形操作、图像处理、地图制图、数据库管理、投影变换等等，它们涉及不同的专业领域。利用构件技术，可以将不同专业领域的具体功能实现为构件，组装形成软件系统，避免了需要同时学习多个专业领域的状况。

图 16-8 给出了一个 GIS 体系结构以及构件的概念模型，在该模型中，体系结构具有多个“插槽”，而构件是“可插入(Plugable)”的，这种类似于计算机主板的结构，通过 DCP 的支持，依照图 16-7 中所描述的工作流结构，可以很容易的实现跨平台的地理信息系统软件。

3．GIS 需求分析

3．1 GIS 需求获取

需求获取是软件开发活动的第一步，获得正确的需求描述是成功软件的前提。一般而言，用户需求分为两类，功能性需求和非功能性需求，前者定义了系统做什么，包括输入、输出

* 关于这点，在地理信息系统标准以及OpenGIS中还将详细论述。

以及其间的转换；后者定义了系统工作时的特性，如效率、可靠性、安全性、可维护性、可移植性等等要求。具体的需求获取内容包括：

1）物理环境，物理设备的位置以及其分布的集中程度；

2）接口，与其它软件系统的接口以及对数据格式的要求；

3）用户或人的因素，包括系统用户熟练程度，使用系统需要接受的训练；

4）功能，系统要完成什么，性能如何；

5）文档，需要哪些文档以及其针对的读者；

6）数据，数据格式、数据精度、数据量、接收和发送数据的频率；

7）资源，使用系统需要的设备，开发需要的人力资源、计算机资源、时间表；

8）安全性，对访问信息的控制程度，数据的备份等；

9）质量保证，对系统的可靠性要求，平均系统出错时间，可移植性，可维护性等。

在地理信息系统需求获取过程中，其内容是基本一致的，只是在数据、组织等方面要进行额外的考虑。

1）数据

数据在一个 GIS 应用系统中，占有举足轻重的位置。进行需求获取时，与数据有关的因素包括，

（1．1）数据的输出样式：包括屏幕显示、Web 发布、出版、工程图等；

（1．2）输出数据的内容和要求：输出数据要包括哪些内容，数据的精度，比例尺等；

（1．3）数据的分布性：数据是集中管理还是分布管理；

（1．4）现有的纸质地图：现有的纸质地图的内容，其比例尺、时效性、是否涉及保密；

（1．5）现有的电子数据：数据形式（栅格/矢量/属性数据库）、数据格式、完整性、精度、投影方式、比例尺等因素。

（1．6）数据录入：数据量大小，输入设备，包括数字化仪、扫描仪，软件的支持程度，进行数据录入的人员数目，能否在预定时间内完成数据录入；

（1．7）数据购买：数据量以及价格；

2）人员

因为应用地理信息系统一般是针对专业领域的，在进行需求获取时不仅要考虑一般用户，而且要听取领域专家的意见，将他们的理论、经验模型化，并在系统中完成。

3）组织

现有的组织机构，有关部门的职责，业务流程，GIS 如何在其业务流程中体现以及因为应用 GIS 所引起的可能的组织机构的变化。

进行需求获取的方式是多种多样的，包括面谈、电话访谈、参观、问卷、获取领域相关资料等。在地理信息系统开发中，由于GIS的概念、功能等还没有被用户深入理解并接受，采用GIS专题报告可以很好地激励用户提出需求。如果时间和资金允许，开发原型系统也可以更好地挖掘用户需求*。

3．2 GIS 需求规约

在需求获取阶段，得到了用自然语言描述的用户需求，但是其中存在着不一致性和二义性，这些问题要通过需求规约解决。目前有许多方法用于支持需求规约，如功能分解方法，

---

* 在软件工程领域，有一句名言，“用户只有看到他不希望的东西，才知道想要什么”，说明了开发原型系统的重要性。

数据流方法，信息模型方法（实体关系模型），面向对象方法。每种方法都有相应的概念体系、符号表现和工具支持。

需求规约的结果是形式化或半形式化的，系统的需求报告，它必须完整地刻画问题域，能够适应需求变化，此外，它必须满足用户、分析人员、设计开发人员进行交流的需要，换言之，需求需求报告中的符号、描述，对于各类相关人员，其意义是一致的。

正如一再强调的，面向对象的分析方法很好地解决了问题空间理解、需求易变性、交流等方面的问题。将通用的面向对象分析方法应用于 GIS，固然可以描述系统需求，但是由于 GIS 更加关注于空间对象和空间模型，在 GIS 中要经常处理一些特定的空间对象类以及特定的空间关系（拓扑关系、方位关系、度量关系等），对基本的面向对象方法进行特化扩展，采用特定的符号来表示这些类以及关系，可以使需求报告的表述更加简捷，便于信息交流。

对 Coad/Yourdon 方法的扩展——GeoOOA（部分，George Kösters）

GeoOOA 区分了基本类以及四个地理类类型，下面是 GeoOOA 的类符号：

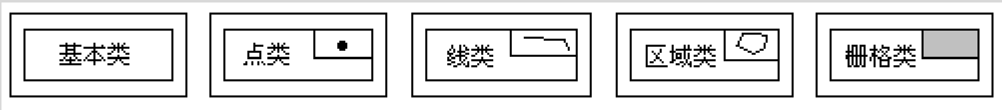


图 16-9：基本类和 GeoOOA 类

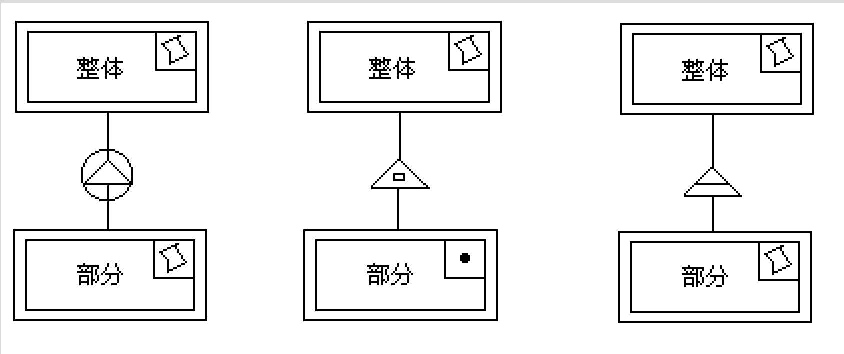
一些服务包括读写操作以及几何体标准服务赋予了 GeoOOA 类，如计算面积，长度，关系检测等等。

地理类形成层次化的一般——特殊关系，栅格类为区域类的特例，而区域，线，点是抽象地理类的特化。

GeoOOA 同时提供了拓扑的整体——部分结构原语，它们只是适用于空间对象，整体部分结构又可以分为：覆盖(Covering)，包含(Containing)，分割(Partition)。

- 覆盖：整体对象和部分对象属于相同的空间类，整体几何体被部分几何体的联合所覆盖；
- 包含：整体几何体包含部分几何体；
- 分割：是一种包含结构，但是有两个特性，即，整体几何体和部分几何体属于同一个类，并且部分几何体形成对整体的分割。

三种原语的图形表示如下：



1）覆盖　　2）包含　　3）分割

图 16-10：GeoOOA 中的整体部分关系

举一个例子说明这三种整体——部分关系，在一个省级行政区（区域类）内，有许多电视台站（点类），每个电视台站有其覆盖范围（区域类），省又分为许多县级区划（区域类），由于台站的覆盖范围之间存在重合，并且一般而言，所有台站的覆盖范围要比省的范围大，所以省行政区和覆盖范围形成覆盖整体——部分关系，而省行政区与台站为包含整体——部分关系，省行政区和县行政区是分割整体——部分关系。

采用 GeoOOA，用图形就直接表述了空间关系，减少了需求报告中文本规约的数量。

4．数据管理设计

数据管理部分设计的目的是确定在数据管理系统中存储和检索数据的基本结构，其原则是要隔离数据管理方案的影响，不管该方案是普通文件、关系数据库、面向对象数据库或者是其它方式的。

目前，主要有三种主要的数据管理方法，即文件、关系和面向对象。

1）普通文件管理：普通文件管理提供基本的文件处理和分类能力；

2）关系型数据库管理系统（RDBMS）：关系型数据库管理系统建立在关系理论的基础上，采用多个表来管理数据，每个表的结构遵循一系列“范式”进行规范化，以减少数据冗余。

3）面向对象的数据库管理系统：面向对象的数据库是一种正在成熟的技术，它通过增加抽象数据类型和继承特性以及一些用来创建和操作类和对象服务，实现对象的持续存储。

不论在分析阶段采用何种方法，已经实现是采用 OOPL 还是非 OOP，都可以选择上述任意的一种方案实现数据的管理。

在地理信息系统软件中，需要管理的数据主要包括：空间几何体数据、时间数据，结构化的非空间属性数据以及非结构化的描述数据。例如，对于地籍管理系统中的地块，

- 空间几何体数据：地块界点的坐标；
- 时间数据：地块存在的时段；
- 非空间属性数据：地块的权属，地价等等；
- 非结构化的描述数据：描述地块所需要的图像、声音数据等等。

为了实现对这些数据的管理，通常的方案包括：

4．1 全部采用文件管理

将所有的数据都存放于一个或者多个文件中，包括结构化的属性数据。采用文件管理数据的优点是灵活，即每个软件厂商可以任意定义自己的文件格式，管理各种数据，这一点在存储需要加密的数据以及非结构化的、不定长的几何体坐标记录时是有帮助的。文件管理的缺点也是显而易见的，就是需要由开发者实现属性数据的更新、查询、检索等操作，而这些，都可以利用关系数据库完成，换言之，利用文件管理增加了属性数据管理的开发量，并且也不利于数据共享。目前，许多 GIS 软件采用文本格式文件进行数据存储，其目的是为了实现数据的转入和转出，与其它应用系统交换数据。

4．2 文件结合关系数据库管理

这是目前大多数 GIS 软件所采用的数据管理方案。考虑到空间数据是非结构化的、不定长的，而且施加于空间数据的操作需要 GIS 软件实现，这样就可以利用文件存储空间数据，而借助于已有的关系数据库管理系统（RDBMS）管理属性数据。采用这种管理方式：

1）空间数据：通过文件进行管理；

2）时间数据：是结构化的，可以利用数据库进行管理；

3）非空间属性数据：利用数据库进行管理；

4）非结构化的描述数据：由于描述数据，不论是文本、图像，还是声音、录象，一般都对应于一个文件，这样可以简单地在关系数据库中记录其文件路径，其优点是关系数据库

数据量小，缺点是文件路径常常会因为文件的删除、移动操作而变得不可靠。如果关系数据库支持二进制数据块字段，也可以利用它来管理文本、图像甚至声音、录象文件。

由于空间几何体坐标数据和属性数据是分开存储管理的，需要定义它们之间的对应关系。通常的解决方案是在文件中，每个地物都有一个唯一标识码(地物 ID)，而在关系数据表结构中，也有一个标识码属性，这样每条记录可以通过该标识码确定与对应地物的连接关系（图 16-11）。

采用该管理方式的缺点在于经常进行根据地物 ID 的查找（既包括从给定地地物查找其对应地记录，又包括根据给定地记录检索相应的地物），使查询、模型运算等一些操作的速度变慢。

地物 ID	坐标
ID1	X1,Y1,X2,Y2 …
ID2	X1,Y1,X2,Y2 …
ID3	X1,Y1,X2,Y2 …
…	…

(a)通过文件管理空间数据

地物 ID	属性 1	属性 2	…
ID1	属性值	属性值	…
ID2	属性值	属性值	…
ID3	属性值	属性值	…
…	…	…	…

(b)通过关系数据库管理属性数据

图 16-11：同时使用文件和关系数据库管理 GIS 数据，
其中利用地物 ID 建立记录之间的连接关系

4．3 全部采用关系数据库管理

在这种管理方式中，不定长的空间几何体坐标数据以二进制数据块的形式被关系数据库管理，换言之，坐标数据被集成到 RDBMS 中，形成空间数据库，其结构如图 16-12。

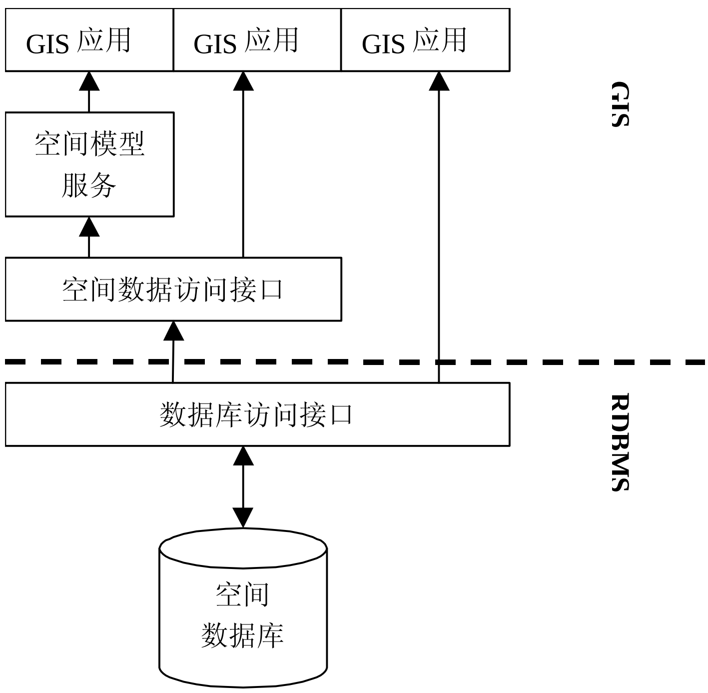


图 16-12：集成化的 GIS 数据管理

可以认为一个地物对应于数据表中的一条记录，这样它带来的最直接的好处是避免了对“连接关系”的查找。目前，关系数据库不论是理论还是工具，都已经成熟，它们提供了一致的访问接口（SQL）以操作分布的海量数据，并且支持多用户并发访问，安全性控制和一致性检查。这些正是构造企业级的地理信息系统所需要的，此外，通用的访问接口也便于实现数据共享。

采用全关系 GIS 数据管理，由于几何体坐标数据不定长，会造成存储效率低下，此外，现有的 SQL 并不支持空间数据检索，需要软件厂商自行开发空间数据访问接口，如果要支持空间数据共享，则要对 SQL 进行扩展。

4．4 采用面向对象数据库(OO-DBMS)管理

如果应用对象数据库管理 GIS 数据，则可以扩充对象数据库中的数据类型以支持空间数据，包括点、线、多边形等几何体，并且允许定义对于这些几何体的基本操作，包括计算距离、检测空间关系，甚至稍微复杂的运算，如缓冲区计算、叠加复合模型等，也可以由对象数据库管理系统“无缝”地支持。

这样，通过对象数据库管理系统，提供了对于各种数据的一致的访问接口以及部分空间模型服务，不仅实现了数据共享，而且空间模型服务也可以共享，使 GIS 软件开发可以将重点放在数据表现以及复杂的专业模型上（图 16-13）。

不过，目前对象数据库管理系统远未成熟，许多的技术问题仍需要进一步的研究。例如，由于支持用户自定义功能，可能会引发对系统的恶意入侵。查询优化也是对象数据库所面临的一个难题，假定要进行查询操作，得到所有有铁路通过并且人口大于 10 万的县，很明显，得到人口大于 10 万的县所需要的计算时间要远远小于得到所有铁路通过的县的时间，系统必须了解这一点，在执行查询时，先得到人口大于 10 万的县，然后再在该集合中依据空间关系进行进一步的查找——而不考虑用户的实际输入次序。该处理过程称为查询优化（Optimize），它使得优化后的计算时间减少。对于这种简单的情况，即结构化查询和空间运算的消耗时间，容易比较判断，但是如果查询语句中包括多个空间运算函数，那么其优化将变得十分困难。

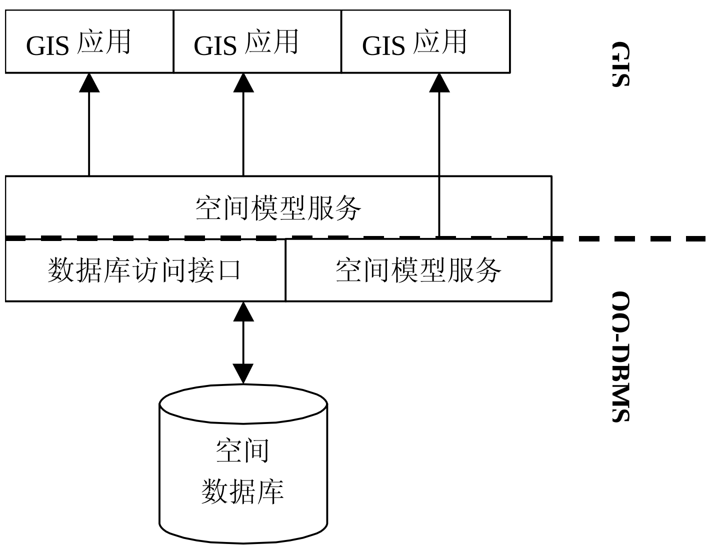


图 16-13：采用 OO-DBMS 进行 GIS 数据管理

5．界面设计

5．1 界面设计原则

在设计阶段，除了设计算法，数据结构等内容外，一个很重要的部分就是系统界面的设计，系统界面是人机交互的接口，包括人如何命令系统以及系统如何向用户提交信息。一个设计良好的用户界面使得用户更容易掌握系统，从而增加用户对系统的接受程度。此外，系统用户界面直接影响了用户在使用系统时的情绪，下面的一些情形无疑会使用户感到厌倦和茫然：

1）过于花哨的界面，使用户难以理解其具体含义，不知从何入手；

2）摸棱两可的提示；

3）长时间（超过十秒）的反应；

4）额外的操作（用户本意是只做这件事情，但是系统除了完成这件事之外，还做了另外的事情）。

与之相反，一个成功的用户界面必然是以用户为中心的，集成的和互动的。

尽管目前图形用户界面（GUI,Graphical User Interface）已经被广泛的采用，并且有很多界面设计工具的支持，但是，由于上述的这些原因，在系统开发过程中应该将界面设计放在相当重要的位置上。

设计用户界面的策略由以下几点构成[Peter Coad,Edward Yourdon]：

1）对人分类

通过仔细研究使用系统的人，对其进行分类，分类的原则包括按照技能层次（初学者，高级人员…），按照组织层次（管理人员，一般员工…），按照身份（职员，顾客…）。通过分类，弄清每类人员使用系统的目的，进而可以确定其相应的人机交互操作。

2）描述人和他们的任务脚本

对人员分类之后，确定每一类人员的特征，包括使用系统的目的，特征（年龄、教育水平、限制等），对系统的期望（必须/想要，喜欢/不喜欢/有偏见），熟练程度，使用系统的任务脚本（scenario）。依据这些特征，可以指导系统的人机交互设计。

3）设计命令层

命令层的设计包括三个方面的工作，即研究现有的用户交互活动的寓意和准则；建立一个初始化的命令层；细化命令层。

在图形用户界面的设计过程中，已经形成了一些形式的或非形式的准则和寓意，如菜单排列（例如，在几乎所有的 MS-Windows 应用系统中，前三个一级菜单项目总是“文件”、“编辑”、“视图”，而最后的两个则是“窗口”，“帮助”），一些操作（例如，打开文件、保存文件、打印）的图形隐喻等等。遵循这些准则，便于用户更快地熟悉系统。

在细化命令层时，需要考虑排列，整体-部分组合，宽度与深度的对比，最小操作步骤等问题。一个层次太“深”命令项目会让用户难以发现，而太多命令项目则使用户难以掌握。

4）设计详细的交互

人机交互的设计有若干准则，包括：

一致性 采用一致术语、一致的步骤和一致的活动；

操作步骤少 使敲击键盘和点按鼠标的次数减到最少；

不要“哑播放” 长时间的操作需要告诉用户进展的状况；

闭包 用一些小步骤引出定义良好的活动，用户应该感觉到他们的活动中闭包的意义；

Undo 人难免做错事，通常在这种情况下系统应该支持恢复原状，或者至少部分的支持。

减少人脑的记忆负担 不应该要求人从一个窗口记忆或者写下一些信息然后在另一个窗口中使用；

学习的时间和效果 为更多的高级特性提供联机参考信息；

趣味和吸引力 人们通常喜欢使用那些感到有趣的软件。

5）继续做原型

通过做原型系统，可以直接的了解用户对设计界面的反应，然后进行改善，使之臻于完美。

6）设计用户界面类

在完成上面的工作后，就可以着手设计用户界面类。在开发 GUI 程序时，通常已经提供了一系列通用界面类，如窗口、按钮、菜单等等，只要从这些类派生特定的子类即可。

7）根据图形用户界面进行设计

目前主要的 GUI 包括 Windows，Macintosh，X-Windows，Motif 等，基于它们开发应用软件可以使界面的设计简单化，但是事先要清楚其特性，如事件处理方式等等。

5．2 GIS 界面设计中的要素

对于成功的 GIS 软件，好的界面是不可或缺的，在进行设计时，同样要遵循上述的原则和步骤。对于地理信息系统软件而言，其界面需要允许用户选择并检索相应的空间数据，操作这些数据，并且表现分析的结果。对于基本的数据检索，操作和表现，与普通的软件是一致的，在 GIS 中要考虑的是以下几个要素：

5．2．1 数据选择

选择数据采用的过滤器可能包括空间的和非空间的属性，或者是两者的结合，例如（针对一个县级行政区划数据）：

- 检索所有人口大于 10 万的县；
- 检索所有有铁路通过的县；
- 检索所有有铁路通过并且人口大于 10 万的县；

用户可以通过输入一个命令语句，通过菜单选择，填充一个表单（Form），或者通过直接的操作——例如使用鼠标来选择数据。利用命令语句需要了解数据表的结构，并且需要对传统的 SQL 进行扩展以支持空间过滤。直接操作选择需要数据显示在屏幕上，它和 SQL 查询的组合通常要间接的完成。当一些地物显示在相近的位置时，直接操作选择会有二义性。为了方便用户操作，要提供一些额外的操作，如漫游、放大显示、缩小显示等。

5．2．2 数据表现

一种好的数据表现形式有利于用户直接操作以进行进一步的分析，与标准的关系数据库相比，在地理信息系统中，需要更多地考虑其图形显示。图形表现的相似性或者差异可以表达给用户这些地物对象具有某种程度相似性，一些图形显示的变量有：

- 多边形轮廓：颜色，灰度，黑白；线型；
- 多边形填充：颜色，灰度，黑白；填充模式；
- 线：颜色，灰度，黑白；线型；

- 符号：颜色，灰度，黑白；形状；大小等等。

对于图形显示而言，需要进行仔细设计，以正确地表达地物对象的含义，并被用户理解。

5．2．3 数据处理

数据处理由一系列空间的和非空间的操作组成，一个设计良好的界面使实现这些操作更加容易。与标准的关系数据库相比，GIS 所管理的数据更具有面向对象的特征，所以一个面向对象界面有利于用户与系统的交互操作，完成数据处理。在 GIS 软件中，面向对象的界面设计包括将地理实体，如点、线、多边形以及一些操作以象形的符号表现出来，而用户可以通过简单的点击、拖放等操作实现相应的数据处理。

下面列出了一些具有图形寓意的操作，很容易将图形表现与对应的操作联系起来。

- 创建：根据给定的图形实体以及相应的属性，创建一个空间对象；
- 删除：删除一个选定的空间对象；
- 集合：根据选择的对象形成一个集合；
- 更新：只显示最后一次操作的结果；
- 叠合：相当于集合操作中的并运算；
- 求交：相当于集合操作中的交运算；
- 求差：相当于集合操作中的差运算；
- 转换：对选择的地物进行比例、移动、镜象、仿射变换等操作；
- 检查点：设置检查点，当对后面工作不满意时，可以回退到该点；
- 回送：返回上一个检查点的状态；
- 提交：将所有的处理结果传给数据库，并更新之。

5．2．4 SQL

传统的 SQL 并不能处理空间查询，这是由于关系数据库技术的弱点造成的，对于 GIS 而言，需要对 SQL 进行扩展。目前正在制定的 SQL/MM，主要应用于多媒体数据，其中包含了全面的 GIS 操作集合。

采用新的 SQL 标准，带来了概念上的改变，对于传统的 SQL，要实现空间操作，需要将 SQL 命令嵌入一种编程语言中，如 C 语言；而新的 SQL 允许用户定义自己的操作，并嵌入到 SQL 命令中。

这种扩展的 SQL 实际上是增加了面向对象的支持，在地理信息系统中，称为 GeoSQL，GeoSQL，可以写出形如：

```
SELECT Soils.Map
From Soils,Parcels
Where Parcels.Value>6000 and Overlay(Soils,Parcels);
```

的查询命令。

由于完全实现 GeoSQL 仍然有相当的难度，所以目前大多数 GIS 软件的实现思路是分别输入标准 SQL 查询语句以及空间查询（图 16-14），然后分开处理，最后将结果合并，其缺点是不够灵活。

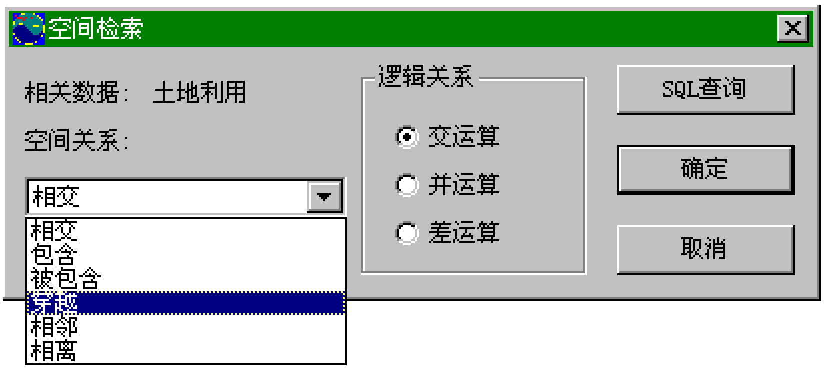


图 16-14：一个实现空间查询的界面，将空间关系和标准 SQL 分开输入

5．2．5 可视化

由于地理信息系统是基于图形的，其分析和解释的结果通常是以可视化的形式表现出来。可视化是指为了识别、沟通和解释模式或结构，概括性地表现信息的过程。空间分析需要考虑信息模式以及空间特征的感受，对于 GIS，可视化可以描述为从信息到知识的转化过程。对于地理信息系统，除了以可视化的形式表现各种信息，实现表达的所见即所得(WYSIWYS, What You See Is What You Get)亦是界面设计的重要原则。

5．3 GIS 界面样式

在 GIS 软件界面设计中，有三种基本的用户界面样式，即基于命令的界面、菜单驱动的界面以及基于工作流的 GUI 界面，这三种界面对于实现和使用各有其长处和短处，在具体实现时，可以同时支持一种或几种样式。

5．3．1 基于命令行的 GIS 界面（图 16-15）

命令行是最简单的界面样式，并且很早就已经在各种操作系统软件中被采用。它只使用文本语言，要求用户了解可以使用的选项，这需要记忆各种命令或者不断查找帮助文档。基于命令行的界面不提供任何提示信息和建议，这使用户要依赖于印刷文档来学习系统。采用命令行界面需要开发一个命令行解释器，在命令行界面软件中，功能模块之间关系较为简单，常常是一个模块的输出作为另一个模块的输入，便于开发实现。利用批命令文件或者脚本文件，可以依次完成多步操作，这是命令行界面的长处。

对于 GIS 软件，因为包含大量的图形操作，所以采用命令行界面时，需要有一个图形窗口以显示操作结果，这样命令行界面起到控制台的作用。由于支持批命令和脚本文件，可以使用命令行界面来实现批量的、流程化的、耗时的数据处理。

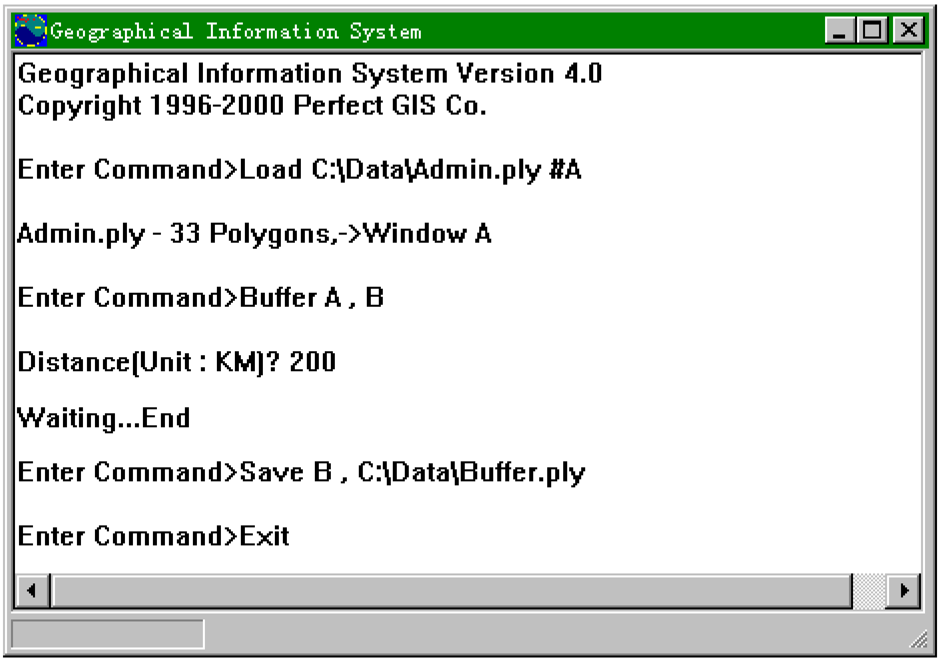


图 16-15：基于命令行的 GIS 界面

5．3．2 菜单驱动的 GIS 界面（图 16-16）

在 MS-Windows 成为 PC 上的主流操作系统之后，菜单驱动的用户界面几乎在所有的应用软件中被采用。它按照层次，列出了系统提供所有操作，用户可以通过键盘或者指点设备，通常是鼠标，来选择并执行一个操作。每个菜单项目都有相应的帮助信息，便于用户随时参看。

菜单驱动界面最大的长处在于界面友好，便于用户掌握系统。但是对于高级用户而言，与命令行界面相比，它往往显得不够灵活而且效率低下。在 GIS 中，往往需要连续地对批量数据进行处理，并且需要较长的计算时间，这种情况下采用菜单界面就变得不可忍受。

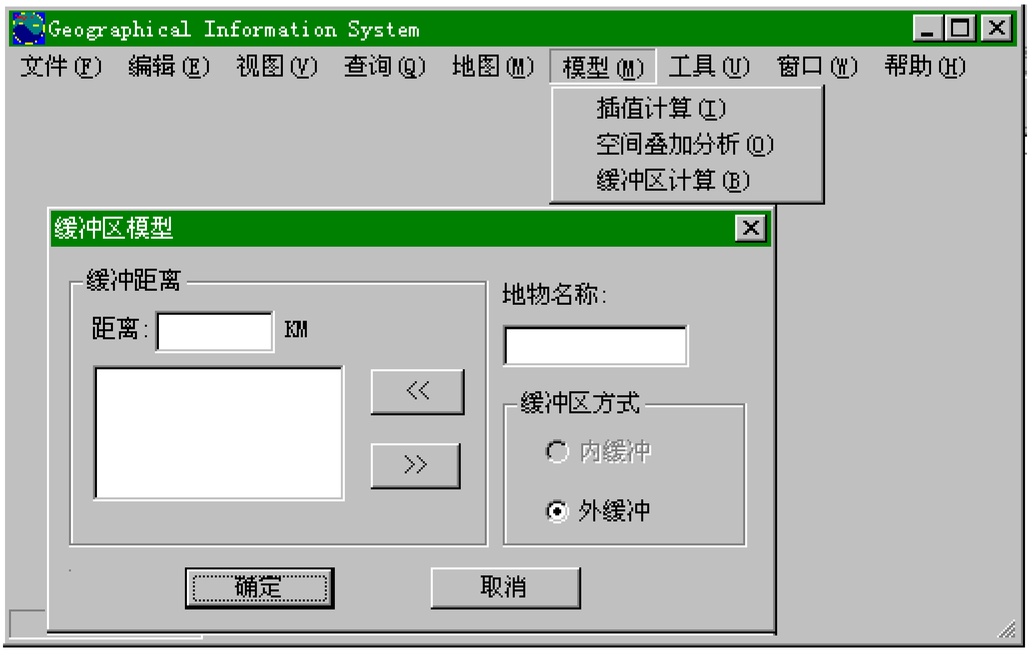


图 16-16：Windows 环境下菜单驱动的 GIS 界面

5．3．3 采用数据流图的 GIS 图形用户界面（图 16-17）

在图形数据流界面中，操作者通过一种“可视化的语言”，而不是严格的文本来控制系统，其中，系统用图形符号来表现其提供的功能，称为“图标”。图标不仅可以表示操作，也可以表示数据或者硬件设备。

在基于数据流图的用户界面中，用户可以通过“拖放（Drag and Drop）”操作来实现相应的操作（例如，可以将表示插值计算的图标拖放到表示等值线数据的图标上，以进行等值线插值运算），设计和组织数据处理流程。

数据流图界面适用于数据流清晰，相对比较简单的系统，如数字图像处理软件；其缺点是实现较为困难。

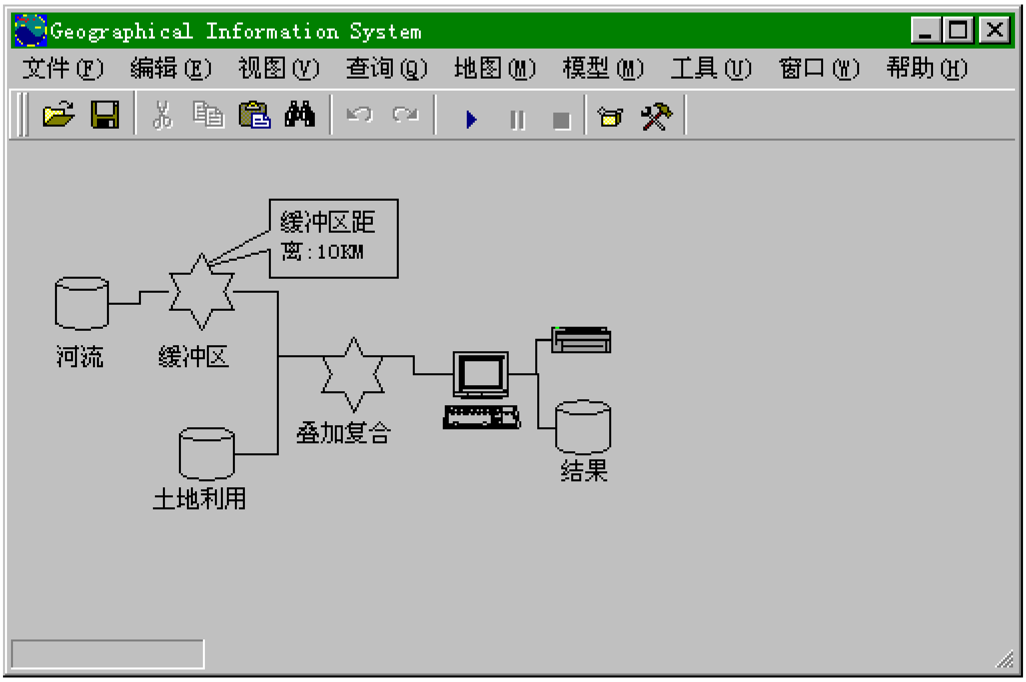


图 16-17：采用数据流图的 GIS 图形用户界面

6．GIS 中的设计模式

随着面向对象技术的广泛应用，软件复用在越来越多的开发过程中被采用。在研究软件复用的过程中，设计模式(Design Pattern)的概念被提了出来。所谓设计模式，简单地理解，是一些设计面向对象的软件的经验总结。正如 Alexander 针对建筑领域所说的：“每个模式描述了一个在我们身边一再发生的问题，它告诉你这个问题的解的关键，以使你可以成千上万次的利用这个解，而不需要再一次去解它。”，在软件开发过程中使用设计模式，可以利用已有的设计经验，指导软件复用。

一个设计模式，一般包括以下四个基本部分：

1）**模式名称**：描述一个设计问题、它的解法和后果；

2）**问题**：告诉什么时候要使用该设计模式，解释问题及其背景；

3）**解决方案**：描述设计的基本要素、它们的关系、各自的任务以及相互之间的合作；

4）**后果**：描述应用设计模式之后的结果和权衡。

E.Gamma 提出了 23 个面向对象的设计模式，这些模式抽象层次较高，可以应用于所有

软件的开发过程。在地理信息系统开发中，经常会遇到本领域的特定的一些问题，并且已经形成了有效的解决方案，对其进行归纳总结，形成相应的设计模式，对于 GIS 软件开发，有着重要的意义。下面遵循 E.Gamma 的格式给出了一个 GIS 设计模式范例。

名称：

过滤和精化（Filter and Refine）

内容：

定义了对一个大数据量的空间数据库的访问接口

另外的名称：

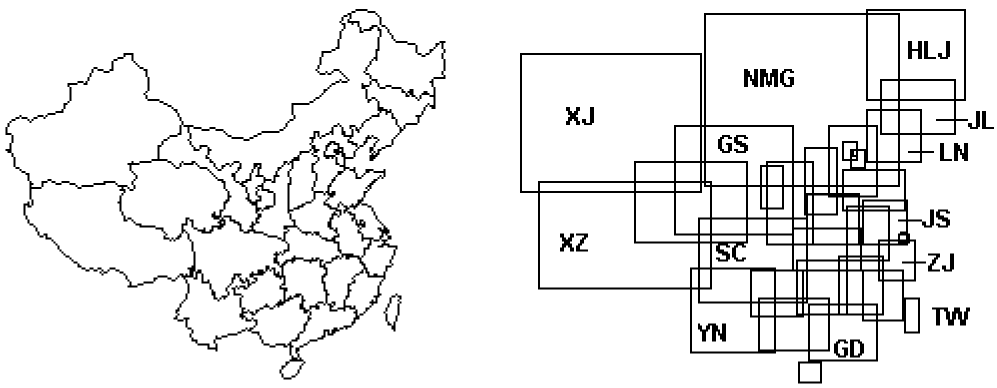
空间索引和最小外包矩形

动机：

在空间数据中，一个通常的查询操作是空间检索，即根据给定的点或者区域得到相应的地物对象，在这些情形中，系统必须要有快的响应速度。

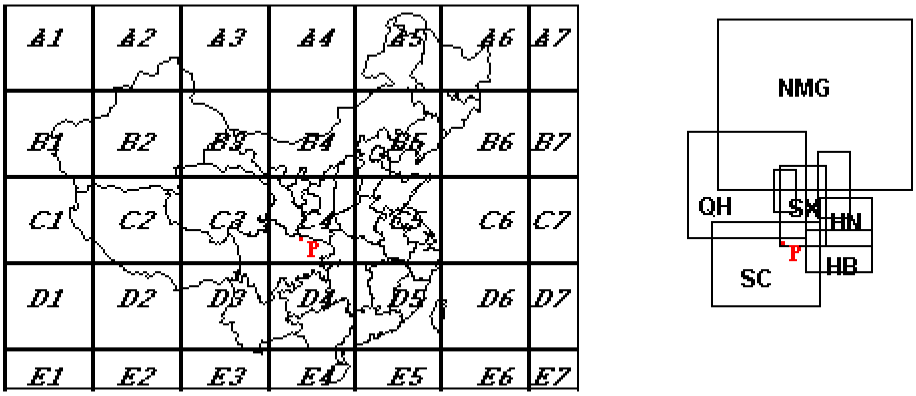
一个最简单的解决办法是依次得到每个地物A_i，并计算A_i和给定的点（或区域）的空间关系，如果符合条件，则返回该地物对象。这个方法的缺点是显而易见的，首先，它需要对空间数据库中的所有地物进行比较运算；其次，空间关系的计算需要大量浮点运算，直接进行给定点（或区域）和地物的空间关系，速度会很慢。

为了提高空间检索效率，有两个解决的途径。首先，可以计算每个地物的最小外包矩形（MBR-Minimum Bounding Rectangle），这样进行空间关系计算时，可以先通过外包矩形来判断，可以排除掉根本不可能具有相交或者包含关系的情形，然后再按照常规的算法（如射线算法等等）进行计算。其次，考虑到采用MBR之后，仍旧要计算每一个地物A_i，当地物数目很多时，依然需要较长的查找时间。解决该问题的一个方案是将数据库的空间范围进行分割，一般是划分成为矩形，然后计算并记录每个矩形包含或者相交的地物，形成空间索引。在进行空间检索时，根据给定的点（或区域）得到其对应的索引块，这样就可以只判断与索引块相关的地物，从而减少了查找时间。通常前一个操作称为精化，后一个操作叫做过滤。下面以中国行政区划数据(图 16-18-a)简单描述该过程。



(a) (b)

图 16-18：中国行政区划和对应的 MBR



（a） (b)

图 16：通过空间索引进行过滤

如图 16-19-a 所示，每一索引块都记录了其相应的地物，形如：

D5→湖北，江西，浙江，台湾，湖南，广东，福建

C4→内蒙古，甘肃，陕西，山西，河南，湖北，四川，青海，宁夏

而每个地物都记录其最小外包矩形，如图 16-18-b 所示。

当根据给定点 P(图 16-19-a)查找包含该点的地物时，首先判断 P 位于 C4 索引块，这样就可以只判断 P 与 C4 索引块对应地物的关系；再由外包矩形，可以得到 P 只可能在四川省或陕西省内(图 16-19-b)，然后就可以采用传统的空间关系计算方法，得到确切的结果——四川省。

应用：

过滤和精化可以应用于根据点或者区域对大数据量的空间数据库的检索（点，PAM-Point Access Method；区域，SAM-Spatial Access Method），而空间数据库由线或多边形地物组成。

结构：

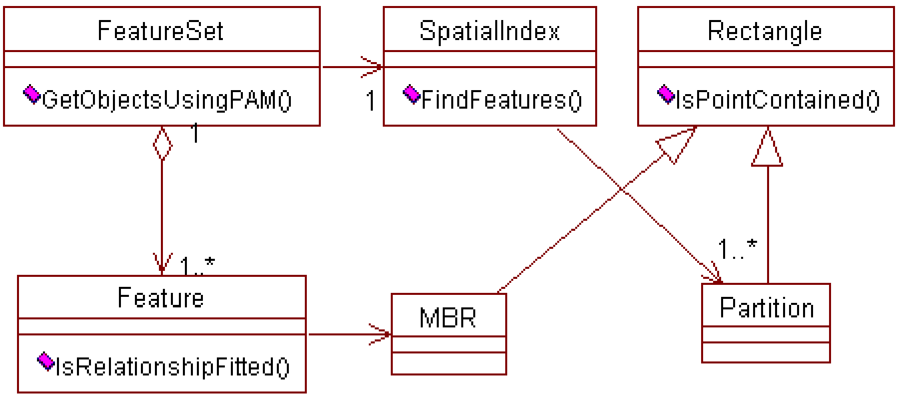


图 16-20：过滤和精化模式结构图

参与者：

- **FeatureSet:**

——定义由地理要素组成空间数据库。

- **Feture:**

——实现了地物对象。

- **SpatialIndex:**

——管理和维护空间索引，可以得到每一索引块对应的地物，提供过滤功能。

- **MBR:**

——地物的最小外包矩形，通过 MBR 可以实现精化功能。

- **Partition:**

——对应于空间数据分割后的每一块矩形区域。

协作：

FeatureSet 利用 SpatialIndex 得到可能符合条件的地物对象集合，然后利用 MBR 去除绝对不可能满足条件的地物。

对空间数据库的任何涉及空间坐标的编辑修改都要重新计算 MBR 和维护空间索引，在要求精度较低的情况下，可以利用最小外包矩形建立空间索引。

后果：

很明显，因为要记录空间索引和最小外包矩形，采用过滤和精化增加了内存开销。索引块的确定是一个重要的因素，如果太小，会有大量地物跨越多于一个的索引块，进一步增加了数据量；而如果太大，过滤的效果不会很明显。如果地物的空间分布不均匀，一般要采用不均匀的分割策略，如四叉树等等。此外，为了保证查询结果的正确性，对地物的每一次编辑修改，都要重新计算其最小外包矩形，并维护空间索引，这需要额外的时间开销，并增加了程序的复杂度。考虑到上述因素，对于数据量较小的情形，一般不建立空间索引。因为点地理要素结构简单，并且数据量小，并且容易计算空间关系，所以对于由点地物对象构成的数据，一般不建立空间索引，并且对于单个点而言，最小外包矩形也是没有意义的。

变化：

为了提高查找效率，可以建立基于四叉树、平衡二叉树等基于树数据结构的空间索引。

可以使用最小外包直角多边形代替最小外包矩形，前者更好地逼近几何实体的形状，提高了精化的效率，但是需要更多的存储空间，并且计算直角多边形运算量较大。

实现：

下面是采用 C++语言描述的采用过滤和精化的空间检索过程，为了简便起见，只是给出了基于点的访问(PAM)，其中 FeatureSet 类表示由地理要素组成的空间数据库，FeatureSet 的成员变量 m_SpatialIndex 为其空间索引，FeatureSet 类的方法 GetObjectsUsingPAM 目的是根据给定的点得到符合条件地物的列表；Feature 类表示地理要素，Feature 类的成员变量 m_rectMBR 为其最小外包矩形。

```
void FeatureSet::GetObjectsUsingPAM(Point ptPoint,Array *arFeatures) const {
    Partition *pPartition = GetRelativePartition(ptPoint);
    Array arRelativeFeatures;
    int nFeaturesCount =
        m_SpatialIndex->FindFeatures(pPartition,&arRelativeFeatures);
    Feature *pFeature;
    for(int i=0 ; i<nFeaturesCount; i++) {
        pFeature = (Feature *)arRelativeFeatures.GetAt(i);
        if(!pFeature->m_rectMBR.IsPointContained(ptPoint))
            continue;
        if(pFeature->IsRelationshipFitted(ptPoint))
            arFeatures->Add(pFeature);
    }
}
```

广义上讲，一个设计模式可以是一个算法、一种数据结构，但是在实践中，模式一般由一组协作的类构成，可以完成特定的任务。在地理信息系统软件开发过程中，利用形式化、半形式化的语言来记录设计模式，可以较好地指导软件复用，提高软件生产率。

7．在 GIS 项目中使用 CASE 工具——空间过程支持与空间数据配置管理

CASE（计算机辅助软件工程，Computer-Aided Software Engineering），是辅助计算机软件开发的计算机技术，包括在软件开发、维护过程中提供计算机辅助支持以及在软件开发、维护过程中引入工程化方法。CASE 工具是一类特殊的软件工具，用于辅助开发、测试、分析和维护另一个计算机程序及相关文档。CASE 工具可以分为以下一些类型，即管理工具、编辑工具、配置管理工具、原型工具、方法支持工具、语言处理工具、程序分析工具、测试工具、调试工具、文档工具和再工程工具等，它们可以用来辅助软件开发过程中不同的活动。

上述的 CASE 工具，同样可以直接应用于 GIS 软件开发过程中，辅助实施相应的软件开发活动。此外，还可以针对 GIS 领域，对 CASE 工具进行特化，使之更好地支持地理信息系统开发活动。下面讲述的在 GIS 开发中，CASE 工具另一个方面的应用：应用配置管理工具实现空间过程支持和空间数据配置管理。

软件配置管理(Configuration Management)是一种标识、组织和控制修改的技术，可以使软件开发过程由于变更修改引起的混乱降低到最小程度。目前，配置管理工具已经比较成熟，提供的基本功能包括：

1）配置标识

在软件开发过程中，各种相关的文档在不断地变更，在变更过程中每一个时刻的内容都称为一个配置，可以命名配置以实现控制和管理。

2）版本控制

版本控制用来管理软件工程过程中所建立起来的配置对象的不同版本。

3）变更控制

变更控制通过“检出（Check out）”和“登入(Check in)”机制实现在多人共同开发软件时，不至于因为共同修改同一文件引起混乱。

4）配置状态报告

配置状态报告通过系统地记录开发过程，反映开发活动的历史情况。

5）配置审核

配置审核的目的是为了证实整个软件生存期中各项产品在技术上和管理上的完整性，并保证所有文档的内容变动不超出当初确定的软件要求范围。

实际上，通过控制不同版本的配置，可以实现整个软件过程的管理，跟踪每份文档的变化，确定文档之间的依赖性。这些特性，同样可以应用于空间过程和空间数据的管理和控制。

在一个较大规模的 GIS 应用软件的建立过程中，常常要处理大量的数据，概括地说，具有以下特征：

1）数据量大；

2）数据经常变更，比如从遥感图像解译得到的土地利用图以及地籍图；

3）对一份或多份数据的空间模型运算，可以得到派生的数据，换言之，数据之间通过空间模型产生了依赖性；

4）空间模型也是在不断变化。

这时，数据变得混乱而难于控制，管理人员无法了解整个项目的进度。在 GIS 中引入配置管理的概念，可以管理空间数据和空间模型的多个版本，进而支持空间过程，辅助 GIS 应用系统的建立。

过程(Process)的概念目前被应用于不同的领域，指被“元程序(Meta-program)”控制的

计算机程序以及数据交换复杂的序列。过程的概念，在解决与异种平台和应用环境相关的一些问题时，是非常有帮助的。这样的问题包括：

1）互操作，它可以在过程的基础上得到解决而不是限于普通的情形。

2）分布，在基于已经存在的工具编写分布应用程序时，提供了充分的支持。

3）前向恢复，通过使用数据库持久记录过程中的每一步来得到保证。

4）监控，基于保存在数据库中的过程的当前状态和前状态来实现。

5）历史回溯，基于对存贮过程状态数据库的查询和数据挖掘工作来实现。

这些问题在许多不同的应用领域很普遍地出现，从虚拟企业和商业环境到软件工程和科学数据管理，这说明了过程概念应用的广泛性。

对于地理信息系统而言，大部分的研究活动集中于以下一些问题：空间数据表现，索引，存储和检索。可以说，这只是试图针对传统数据集进行数据库技术扩展的结果。而空间过程问题，不仅仅只是与空间模型有关，下面用一个例子进行说明。

图 16-21 是一个典型的空间模型组合，形成应用于不同数据集的变换序列*。例如，高程采样数据作为数值高程重建算法的输入，生成带有内插高程的地形图。然后地形图被坡面分析程序使用，以提取不同坡面的坡长，坡度和坡向。这些结果，连同该区域的土壤抽样数据和植被覆盖信息，被用于作为试图预测该区域未来侵蚀模式模型的输入数据。植被覆盖数据通过综合分析土壤抽样数据，植被抽样数据和卫星图像得到。植被覆盖数据也作为植被演替模型的输入数据，该模型根据预测的侵蚀数据，估算未来可能的植被覆盖变化。在考虑到降水量的影响时，也是相似的逐步的过程。图 21 表现了所有的空间模型组织在一起，形成一个复杂的操作集合，表现不同的地理现象。

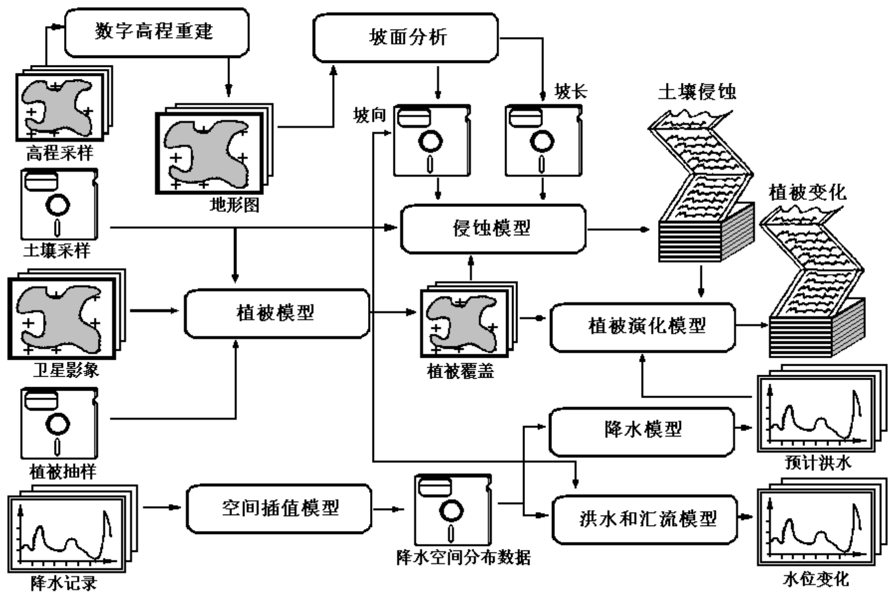


图 16-21：空间过程的例子：水土流失模型[G.Alonso]

这个例子阐明了一些需求，很自然，这些需求与过程管理的需求非常相似。下面列出这些必要的、简化的需求，包括模型语言，分布和并行以及查询能力。

1）模型语言：在任何一个过程支持系统中，其关键成分就是可以表述过程的语言。给定一个复杂的地理过程模型，模型语言必须是结构化的，允许嵌套，并且可以被方便地重用。任何地理过程应该可以方便地被用于构造一个更大的地理过程。此外，考虑到复杂的执行环境，该语言必须提供对事件和例外处理机制的支持。事件可以被用于在数据集发生改变时，

---

* 下面，空间模型和空间过程都被使用，抽象的看，模型是地理现象的表现，而过程是用程序和数据集描述地理现象。

通知系统并且触发执行模型，以产生更新的数据。类似地，考虑到例外，必须存在一个可靠的机制来处理它们，以避免系统偏离预先描述的行为，造成整个模型执行的退出。由于在工作流系统中，该语言必须允许定义和注册外部对象以及应用。在地理过程的应用中，更应该这样，这是因为在地理过程管理中，不论算法还是空间数据都将外在于系统，外部实体的注册是互操作性问题的基础。

2）分布和并行中的前向恢复：为了降低这种复杂的地理过程执行的开销，只要可能，它们的不同的步骤应该并行。一般来讲，我们假定基本的平台是一组(cluster)微机或工作站，不要求是相同的操作系统。地理过程中的每一个步骤都可以被指派到这一组中的不同的结点上，因此采用并行机制是模型固有的特征。同样地，多处理器的机器也采用类似的实现。作为其直接后果，并且由于模型的成本和执行它们所用的时间，必须有一种机制，以避免当发生错误时，所有的计算丢失。这就是前向恢复的概念，它可以从由于错误引起的中断处恢复执行。如果提供一个系统，其中复杂的地理过程只能从头到尾的执行而不能在中间中断，也不能动态修改地理过程以纠正错误，这样的系统将是无用的。通常的，构造一个地理过程与其说是为了得到其结果，不如说是测试相关地理模型的合法性和可用性。基于这些理由，系统必须支持单步执行，并且能够在任何执行点上停止，进行检查，作出改变以及恢复执行。这样的功能只能通过对地理过程执行的精确监测来实现。

3）依赖性查询能力：在地理过程中，如果对模型和输入数据不了解，那么就不能很好地解释利用该数据和模型产生的数据。这样就引出了那些众所周知的问题如“族系跟踪(Lineage tracking)”，“变化传播(Change propagation)”以及“版本”，这些概念与更加一般的过程中的历史跟踪没有什么不同。在这种环境中，将有这样一些典型的查询问题，例如“哪一个模型用到了算法 X？”，“”如果数据集 Y 发生改变，将会产生什么结果？”以及“哪些数据用于产生数据 Z?”。因此，系统必须支持如下的一些功能，包括：自动变化传播（当输入数据发生改变时，重新执行模型，以创建新版本的输出数据)，变化通知以及过程控制（只要一个数据集存在，创建它所使用的模型将保留一个拷贝；同样的，只要有其它过程使用了一个子过程，也保留该子过程的拷贝），这样系统才能真正有用。只有具有合适的机制来跟踪数据之间的依赖性，并且有高效的方法从这些依赖性中提取信息，系统才能很好地提供上述功能。

过程建模是当前软件工程一个重要的研究方向，通过形式化地描述开发过程，进而实现更加“精确”的过程管理，它同样可以应用于空间过程。由于过程本身的复杂度，使得难以建立完整的过程模型，利用配置管理工具，可以在一定程度上实现数据、模型乃至过程的控制管理。

在信息技术产业，要害一个人，就让他去开发一个市场标准。

虞有澄（前 Intel 副总裁）

第十七章 地理信息系统标准

导读：信息技术的标准对于产业的发展有着重要的意义，GIS 也不例外，目前 GIS 标准主要集中于空间数据模型和空间服务模型以及相关领域，此外还包括流程、认证等等。本章介绍了两个主要的 GIS 标准：ISO/TC 211 和 OpenGIS。

1. 地理信息系统标准简介

随着 GIS 技术的发展，特别是网络技术应用到地理信息系统建设中，与它有关的标准化也成为一个必须解决的问题。一个好的标准是促进、指导和保证高效率、高质量地理信息交流不可缺少的部分。

在信息技术领域，标准和规范按照其使用状态，可以分为两种，即实际使用的标准和法律意义上的标准。前者是在不断的实践过程中，有关机构、团体和组织自发达成的被广泛接受的标准，如 TCP/IP 协议，OpenGIS 规范；后者通常是为了政策或管理的目的，通过法律制定的标准，如 FGDC 制定的空间元数据内容标准。

按照管辖地区的大小，制定信息技术的标准化组织可以分为五个层次（Cargill），即国际级标准化组织，如 ISO；区域级标准化组织；国家级标准化组织，如美国国家标准化组织 ANSI 以及美国联邦地理数据委员会；政府和用户级标准化组织，在 GIS 领域，OGC(OpenGIS) 就属于该层次；以及补充性标准化组织。

目前在中国，GB 系列中与 GIS 有关的标准主要是一些地理编码标准，包括：GB2260-80《中华人民共和国行政区划代码》，GB/T13923-92《国土基础信息数据分类与代码》，GB14804-93《1：500、1：1000、1：2000 地形图要素分类与代码》，GB/T5660-1995《1：5000、1：10000、1：25000、1：100000 地形图要素分类与代码》等。

通常，信息技术的标准和规范可以分为以下五个方面：

1）硬件设备的标准，在网络技术中，存在着大量这种标准，如 IEEE 802 系列；

2）软件方面的标准，包括操作系统，查询语言，程序设计语言，图形用户界面等等，如 SQL，DCOM，CORBA 等等；

3）数据和格式的标准，包括数据模型，数据库的构建，数据质量和可靠性，地理要素的分类系统，数据格式转换等，在地理信息应用中，空间数据编码规范、元数据标准等就属于该范畴；

4）数据集标准，数据存放的文件格式标准，如美国人口普查局的 TIGER 文件标准等；

5）过程标准，如 ISO9000 系列和 CMM 等，主要是针对系统开发过程的指导。

地理信息系统标准化主要包括后四个方面的标准，具体内容有：软件工具，如文档，设计、验收、评测标准以及软件的接口规范等；数据，包括数据模型，数据质量，数据产品，数据交换，数据显示，空间坐标投影等；系统开发，系统设计过程，数据工艺流程，标准建库流程等；其它，包括名词术语，管理办法等。一般而言，软件工具，系统开发，管理办法等方面的标准可以借用更为通用的信息技术标准规范，所以 GIS 标准主要集中于空间数据以及相关的一系列规范。

应用地理信息系统标准，可以建立一套较为规范数据的录入处理流程，提高工作效率和质量，同时采用一致的数据格式以及空间数据可视化方式，指导数据的使用。

总而言之，在地理信息系统中引入一系列标准，有利于保障地理信息系统技术及其应用的规范化发展，指导地理信息系统相关的实践活动，拓展地理信息系统的应用领域，从而实现地理信息系统的社会及经济价值。基于地理信息系统标准，可以实现不同应用领域地理信息的共享和互操作，这也正是实现数字地球的关键技术之一（图 17-1）。

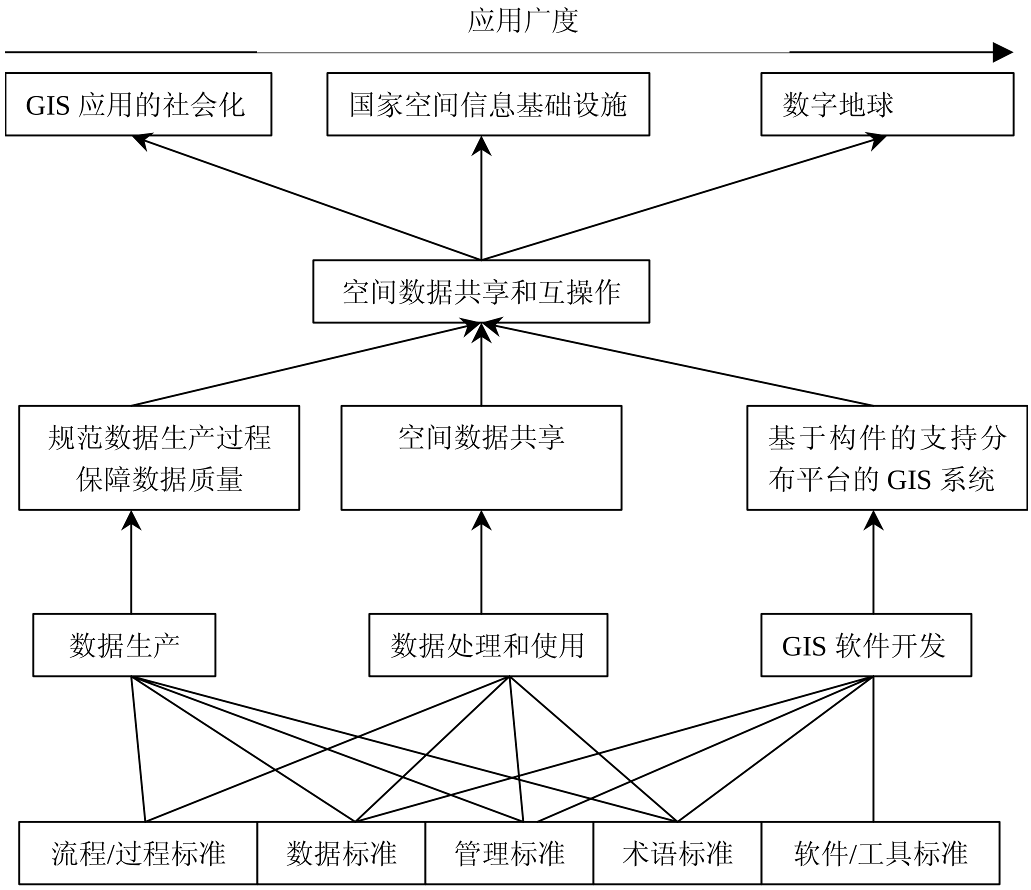


图 17-1：地理信息标准对地理信息系统应用的意义

ISO 的标准制定过程：[Cargill]

ISO 标准的制定过程主要分为以下几个步骤：

（1） 由子委员会起草新的工作内容计划书，提交一个技术委员会投票表决；

（2） 表决通过后，再交由该子委员会负责有关标准制定的各项活动；

（3） 子委员会下属的工作小组起草标准工作草案，交由子委员会投票表决；

（4） 若子委员会统一标准工作草案，该草案成为标准草案计划书，并在整个技术委员会中传阅；

（5） 若技术委员会对草案计划书也能够达成共识，该标准草案将被作为国际标准草案送至技术委员会负责人手中；

（6） 技术委员会负责人将对国际标准草案进行核实，审查是否符合 ISO 的原则和要求；

（7） 若审查通过，该国际标准草案将被送至 ISO 的各国成员中传阅，征求意见，并表决，如果由 75%的成员表决通过，则将提交至 ISO 的执行委员会作为 ISO 的标准出版发行。

目前，有很多个组织和政府部门召集或主持制定地理信息系统标准，其内容主要包括空间数据模型和空间服务模型以及在此基础上的空间数据共享（Data Sharing）和互操作（Data Interoperation）。其中比较主要的是地理信息/地球信息科学(Geographic information/Geomatics)专业委员会制定的 ISO/TC211 地理信息标准以及 OGC(Open GIS Consortium)制定的 Open GIS。

与 Open GIS 相比，ISO/TC211 更为全面，更注重于标准本身的定义，可以指导地理信息系统开发和使用的各个方面；而 OpenGIS 由于有许多著名的 GIS 软件开发商参与，因而更加注重软件的实现。

地理数据的互操作：

随着地理信息系统的发展，需要实现地理数据共享，而互操作则是数据共享的必然产物。

在传统的、以数据集中式管理为特色的地理信息系统应用中，实现数据共享的手段主要是通过数据交换来完成，具体方式包括通过相关表或转换器的直接转换以及基于空间数据转换标准的间接转换。

空间数据的互操作针对异构的数据库和平台，实现数据处理的互操作，与数据转换相比，它是“动态”的数据共享，独立于平台，具有高度的抽象性，是空间数据共享的发展方向。

空间数据的互操作存在着多个层次，从最底层的面向硬件的互操作，到应用层次的信息团体之间的语义共享。下表给出了互操作的层次以及其支持。

表 17-1：地理数据互操作以及其支持（陈述彭）

层次结构		支持
企业	信息	立法、政策、规范、标准
应用	语义	标准、语义数据模型、规程
数据	地理信息系统	DBMS、标准、规程
	数据库	
技术	软件和网络协议	规程、协议、标准
	硬件和网络	标准等

2．ISO/TC211 地理信息标准

ISO/TC 211 地理信息/地球信息科学专业委员会成立于 1994 年 3 月，其目的是为了促进全球地理信息资源的开发、利用和共享，即制定ISO/TC 211 地理信息/地球信息科学标准，它是对与地球上位置直接或间接有关的物体或现象信息的结构化标准。该标准共分为 25 个部分*（截至 2000 年 5 月），主要针对地理信息的内容和相关的方法；各种数据管理的工具和服务及有关的请求、处理、分析、获取、表达；以及在不同的用户、系统平台和位置上进行数据的转换。

1）参考模型(Reference Model)

描述地理信息系统标准的使用环境、使用的基本原则和标准的改造框架，同时也定义了该标准的所有的概念和要素。参考模型是一个独立于任何应用、方法和技术的模型，也是整个 ISO/TC 211 的委员会的工作指南。

2）综述(Overview)

整个 ISO/TC 211 标准系列的介绍和回顾。ISO/TC 211 将是一个完整全面的地理信息系统的标准族，该部分提供给潜在用户一个整体的标准系列和个别标准的综合介绍，包括标准的目的、标准以及标准之间的关系等，使用户可以快速查询到所需要的内容，提高标准的可理解性和可接收性。

3）概念化模式语言(Conceptual Schema Language)

---

* 这 25 个部分主要由五个工作组负责。第一组，由美国召集，负责框架和参考模型，即 1-5,21,24 部分；第二组，由澳大利亚召集，负责空间数据模型和算子，即 7、8、9 和 23 部分；第三组，由英国召集，负责地理空间数据管理，即 10-15 部分；第四组，由挪威召集，负责地理空间数据服务，即 16-19,24 部分；第五组，由加拿大召集，负责第 6,20 部分，第 22 部分由加拿大负责的另一个小组制定。其中后五个的部分的工作刚刚开展。ISO/TC211 委员会同时有一个质量控制小组和咨询小组支持整个标准制定工作。

使用一种标准化的模式语言来促进互操作标准的开发，并提供一个快速建立地理信息标准的基础。这种标准语言是在现有的标准概念化语言之上发展而成的。

4）术语定义(Terminology)

定义了所有 ISO/TC 211 标准中使用的专有词汇，其目的是产生通用的与地理信息标准有关的词汇，供地理信息系统的标准制定者、使用者和开发者使用。

5）一致性和测试(Conformance and Testing)

为了保证所有 ISO/TC 211 标准的一致性而制定的测试框架、概念和方法。建立测试方法的标准和保持一致性的原则可以使 GIS 软件的开发者来核实各类标准的一致性。

6）专用标准(Profiles)

定义所有 ISO/TC 211 标准的子集产品，它确定了在 ISO/TC 211 制定的全部标准的基础上，针对某些具体应用提取出专用标准子集的方法和参考手册。

在 ISO/TC 211 中，定义和描述了一系列地理信息以及地理数据管理和地理过程的标准。其中，某个方面可能有多个标准，如测量标准和编码标准；其它一些标准可能描述了一系列内容，如空间模式标准。在实际应用中，可能只采用某个标准或标准的一部分，甚至是对某个标准进行特化，专用标准给出了使用的指导。

7）空间模式(Spatial Shema)

定义对象空间特征的概念模式，主要从几何体和拓扑关系的角度来制定概念模式。几何体和拓扑关系是地理信息的两个主要特征，它们的标准制定将为其它空间特征标准制定提供方便，同时可以帮助 GIS 开发者和使用者理解空间数据结构。

8）时间尺度子模式(Temporal Subshema)

定义空间实体时间尺度特征的概念，地理信息并不局限于三维尺度，许多地理信息系统需要时间特征。

9）应用模式规则(Rules for Application Schema)

定义地理信息应用的模式，包括地理对象的分类和它们与应用模式之间关系的原则。采用一致的形式定义应用模式，将增强应用之间数据共享能力，并且允许应用之间实时地交互操作。

10）要素分类方法(Feature cataloguing methodology)

定义了对地理对象、属性和关系进行分类的方法论，并且确定建立一个国际化的多语言的分类的可能性。地理信息的类别一般都决定于应用模式，提供一致的分类方法学增强了从一个类别映射到另一个类别的可能性。

11）坐标空间参照系统(Spatial referencing by coordinates)

定义了坐标空间参照系统的概念化模式以及描述大地参照系统的指导，其中也包括一些国际上使用的参考系统，制定坐标空间参照系统同样有助于各类应用之间的交流和数据共享。

12）基于地理标识的间接参考系统(Spatial referencing by geographic identifiers)

定义了间接的空间参照系统的概念化模式。ISO 认为越来越多的有关地理信息的应用使用非坐标类型的参照系统，即间接空间参照系统，例如地址数据，因而有必要产生一套间接参考系统的标准模式。

13）质量原则(Quality principles)

定义了应用于地理数据的质量模式。对地理信息的创建者和使用者而言，质量信息都是十分重要的。一致的质量标准模式，便于一个应用中创建的数据在另一个应用中被适当的评估和使用。

14）质量评价过程(Quality Evaluation Procedures)

给出了对数据质量进行评估和描述方法的指导。关于地理数据质量的评价信息不仅需要

一致的标准，而且需要一个一致的、标准的评估和描述方法。一个标准的评估准则可以保证不同数据集合的质量具有可比性。

15）元数据(Metadata)

定义地理信息和服务的描述性信息的标准。该标准制定的目的是为了产生一个地理元数据的内容及有关标准。这些内容包括地理数据的现势性、精度、数据内容、属性内容、来源、覆盖地区以及对各类应用的适应性如何等。对地理数据进行标准的描述可以使地理信息用户方便地得到适用的数据。

16）空间信息定位服务(Positioning Service)

定义了定位系统的标准接口协议。全球定位系统的发展使得一个地理对象在全球范围内的定位成为可能，定位信息标准接口的制定将促进这些定位信息在各类应用中更有效的使用。

17）地理信息描述(Portrayal)

定义了地理信息描绘方法，不同应用系统之间采用一致的符号表现方法，将便于人们更好的理解和识别各类地理信息。

18）编码(Encoding)

选择与地理信息使用的概念模式相匹配的编码规则，并且定义了概念模式语言之间以及编码规则之间的映射方式。编码规则使得地理信息在以数字形式进行存储和传输时，按照一定的编码语言和系统进行编码。

19）服务(Service)

识别和定义地理信息的服务接口以及与开放系统环境(Open System Environment)模型之间的关系。服务接口的定义有助于不同层次的各种应用访问和使用地理信息。

20）功能标准(Functional standards)

定义了地理信息科学领域已经识别出的功能标准的分类方法。功能标准的分类有利于ISO/TC211 与其它标准的协调一致。标准子集的制定也与功能标准的识别有关。

21）图像和栅格数据(Imagery and gridded data)

为了使 ISO/TC211 能够处理地理信息场模型中的图像和栅格数据，ISO/TC211 需要定义图像和栅格数据标准。它确定了其它组织以及 ISO 其它委员会定义的图像标准，这些标准支持地理信息中栅格和矩阵数据标准的建立。由于地理信息中图像和栅格数据产品的增加，需要该方面的标准的制定。

22）职员的资格认证(Qualifications and Certification of Personnel)

描述了地理信息科学/地球信息学中人员的资格认证体系，定义了地理信息科学/地球信息学与其它相关学科以及专业的边界。详细说明了属于地理信息科学/地球信息学领域的技术。建立了该领域中技术人员、专业人员以及管理人员的能力范围和资格水平体系。

23）覆盖几何和功能的模式(Schema for coverage geometry and functions)

定义了描述覆盖的空间特征的标准概念模式。覆盖通常包括栅格数据、不规则三角网、点覆盖和多边形覆盖。在大量地理信息应用领域中，覆盖是主要的数据结构，包括遥感、气象、地形、土壤、植被等等。覆盖几何和功能的模式将有助于提高地理信息在这些领域内的共享能力。

24）图像和栅格数据的成分(Imagery and gridded data components)

给出了描述和表现图像和栅格数据的概念标准，这包括针对图像和栅格数据的新的工作：应用模式规则、质量原则、质量评价过程、空间参照系统、可视化和服务等，并表明的新的工作与已有的针对矢量数据标准的不同之处。

25）简单要素的访问——SQL 选项(Simple feature access - SQL option)

该部分面向 SQL 环境的简单要素访问实现规范，该实现规范将支持要素的存储、检索、

查询和更新操作。

ISO TC211 标准的各个部分之间具有依赖关系。下表描述了这种依赖性。其中单元格内容为"D"表示相应行的标准依赖于相应列的标准，如“质量评定过程”依赖于“质量原则”，"I"表示有关，"X"表示无关，"C"表示关系尚不明确。

表 17-2：ISO/TC 211 标准前 20 个部分之间的关系

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1:参考模型		I	I	I	X	I	I	I	I	I	X	C	I	X	I	I	I	I	I	X
2:综述	D		D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
3:概念化模式语言	D	I		I	I	I	I	X	I	I	I	I	I	I	I	I	I	I	I	I
4:术语定义	D	D	D		D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
5:一致性和测试	D	I	X	I		D	D	D	D	D	D	C	C	D	D	I	I	I	I	D
6:专用标准	D	I	D	C	I		D	D	D	D	D	C	C	D	D	D	I	D	D	D
7:空间子模式	D	I	D	I	D	I		X	D	I	I	I	I	I	I	X	I	X	I	I
8:时间子模式	I	I	D	I	I	X	X		C	C	X	X	I	I	X	X	I	I	X	X
9:应用模式规则	D	I	D	I	X	I	D	D		D	X	C	D	X	D	X	I	I	I	D
10:地理信息分类	D	I	I	I	X	X	X	I	D		X	X	C	X	D	X	X	X	X	X
11:坐标参考系统	D	I	D	I	D	I	I	X	I	X		X	C	X	X	X	X	X	X	O
12:间接参考系统	D	I	D	I	I	X	I	X	X	X	I		I	X	I	X	X	X	X	X
13:质量原则	C	I	D	I	I	X	I	I	I	I	I	I		I	I	X	X	I	X	X
14:质量评定过程	D	I	D	I	X	X	I	I	I	I	I	I	D		I	X	X	X	X	X
15:元数据	D	I	D	I	I	X	D	D	C	D	D	D	D	D		I	I	D	D	D
16:定位服务	D	I	D	I	D	X	I	I	X	X	D	X	C	X	I		X	X	D	X
17:地理信息描述	D	C	D	I	D	I	D	D	D	I	X	I	I	I	D	X		X	I	X
18:编码	D	I	D	I	X	I	X	X	D	D	X	X	C	X	D	I	I		I	X
19:服务	D	I	D	I	D	I	I	I	I	X	I	X	X	I	D	D	I	X		D
20:功能标准	D	I	D	I	X	C	D	D	D	X	D	D	X	X	D	X	X	X	D	

3．开放的地理数据互操作规范——OpenGIS

OpenGIS(Open Geodata Interoperation Specification,OGIS-开放的地理数据互操作规范)由美国 OGC(OpenGIS 协会，OpenGIS Consortium)提出。OGC 是一个非赢利性组织，目的是促进采用新的技术和商业方式来提高地理信息处理的互操作性(Interoperablity)，OGC 会员主要包括 GIS 相关的计算机硬件和软件制造商（包括 ESRI, Intergraph,MapInfo 等知名 GIS 软件开发商），数据生产商以及一些高等院校，政府部门等，其技术委员会负责具体标准的制定工作。

OpenGIS 的目标是，制定一个规范，使得应用系统开发者可以在单一的环境和单一的工作流中，使用分布于网上的任何地理数据和地理处理。它致力于消除地理信息应用（如地理信息系统，遥感，土地信息系统，自动制图/设施管理(AM/FM)系统）之间以及地理应用与其它信息技术应用之间的藩篱，建立一个无“边界”的、分布的、基于构件的地理数据互操作环境，与传统的地理信息处理技术相比，基于该规范的 GIS 软件将具有很好的可扩展性、可升级性、可移植性、开放性、互操作性和易用性。

OpenGIS 规范主要定义了以下三个模型：

1）开放的地理数据(Open Geodata)模型：

定义了一个概括的、公用的基本地理信息类型集合，该集合可以被应用于特定领域的地理数据建模。OpenGIS 将现实世界抽象成为两类基本对象：要素（Feature）和覆盖（Coverage），前者描述现实世界中的实体对象，后者描述现实世界中的现象。对于要素，将与空间坐标相关的属性抽取出来，称为几何体(Geometry)。同时，OpenGIS 又定义了要素的时空参照系统、语义(Semantics)以及元数据来对要素进行描述，以便于共享和互操作。

2）OpenGIS 服务

定义了一个服务的集合，该集合用于访问地理数据模型中定义的地理类型，提供了同一信息团体（Information Community）内不同用户之间，或者不同信息团体之间的地理数据共享能力。

服务模型中的主要组成为：

（2．1）要素实例(Feature Instance)的创建过程

该过程用到了两个概念，要素模式(Feature Schema)和要素注册(Feature Registry)，前者定义了要素的属性集，包括几何体、描述数据等；要素注册存放要素模式，所有的要使用和共享的要素都要进行注册，要素注册起到了要素"工厂（Factory）"的作用，通过它可以创建要素实例。

这种面向对象的要素实例创建过程便于实现数据的共享，同时又保证信息的封装性。

（2．2）获取地理数据的方法

在 OpenGIS 规范中，建立了一种树状的目录索引结构，其最小单元是要素，通过该目录，可以得到所需要的地理数据。数据的获取在平台间是透明的，即可以跨平台访问数据。

（2．3）时空参照系统的获取和转换

在一个信息团体中是通过时空参照系来转换和解释几何体的，时空参照系统必须能够按照一种统一的标准来定义，并且通过某种机制能够使用这些定义。OpenGIS 描述了注册时空参照系统的机制以及在不同的时空参照系统之间进行地理要素转换的机制。

（2．4）语义转换

不同信息团体之间通过一个语义转换注册器来实现要素语义的转换，注册器包括源要素和要素模式、目标要素和要素模式、以及要素转换器等，可以根据不同的要素转换要求查找到匹配的转换器以进行语义转换。

3）信息团体模型

信息团体模型的目的是，建立一种途径，使得信息团体或用户维护对数据进行分类和共享所遵循的定义；实现一种有效的、更为精确的方式，使不同信息团体之间可以共享数据，尽管他们并不熟悉对方的地理要素定义。信息团体模型定义了一种转换模式，使得不同信息团体的"地理要素辞典"可以自动"翻译"。

OpenGIS 规范包括抽象（Abstract）规范，实现（Implementation）规范以及具体领域（Specific Domain）的互操作性问题，其中抽象规范是 OpenGIS 的基础，也是 OpenGIS 的主体；实现规范定义了抽象规范在不同分布计算平台上的实现，目前 OGC 已经定义了针对 CORBA，OLE/COM 和 SQL 的简单要素访问的实现规范；针对领域的互操作性研究通过提取领域的互操作性用例(Use case)，检验抽象规范能否满足该领域的需求，它是抽象规范的扩展。

抽象规范建立了一个概念模型，并将其文档化，采用了在面向对象技术中通用的 UML 作为其形式化的建模语言。抽象规范通过对现实世界的描述，建立了系统实现与现实世界之间的概念化的联系，它是与具体的软件实现无关的，而只是定义了软件应该实现的内容。

目前，抽象规范共分为 17 个主题*：（截至 2000 年 1 月）

- 综述(Overview)
- 要素几何体(Feature Geometry)
- 空间参照系统(Spatial Reference Systems)
- 位置几何体结构(Locational Geometry Structures)
- 存储功能和插值(Stored Functions and Interpolation)

---

* 这些主题类同于ISO/TC211 中的对应部分，并且一些概念，如"实体"、"覆盖"、"实体之间的关系"等，在第二篇已经详细描述，在此不再赘述。

- 要素 (Features)
- 覆盖类型及其子类型(The Coverage Type and its Subtype)
- 地球影象(Earth Imagery Case)
- 要素之间的关系(Relations Between Features)
- 质量(Quality)
- 要素集合(Feature Collections)
- 元数据(Metadata)
- OpenGIS 服务体系结构(OpenGIS Service Architecture)
- 目录服务(Catalogs Service)
- 语义和信息团体(Semantics and Information Communities)
- 图像使用服务(Image Exploitation Service)
- 图像坐标转换服务(Image Coordinate Transformation Service)

抽象规范的十七个主题之间同样具有相互的依赖性，图 17-2 描述了这种依赖性。

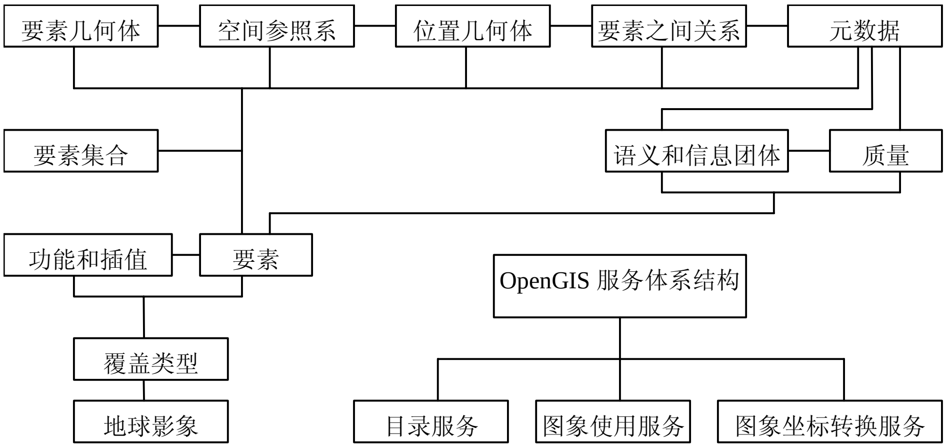


图 17-2：OpenGIS 抽象规范各个主题之间的依赖关系[OGC]

OpenGIS 的实现需要分布计算平台(DCP-Distributed Computing Platform)的支持，图 17-3 描述了遵循 OpenGIS 规范，地理信息处理在不同 DCP“场景”中的实现。但是抽象规范本身与 DCP 无关，它可以在任何 DCP 上实现。

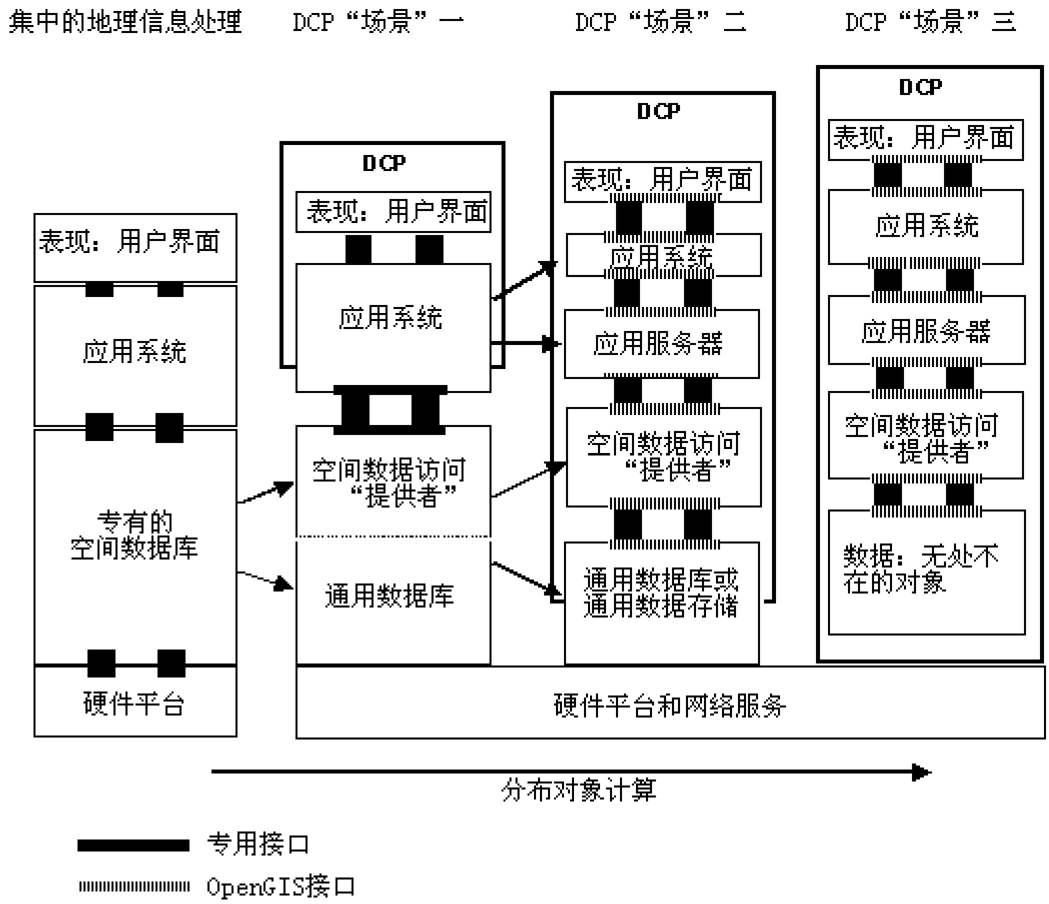


图 17-3：遵循 OpenGIS 规范，地理信息处理的在不同分布计算“场景”上实现[OGC]

其它有关地理信息系统标准制定组织和标准

1）美国联邦地理数据委员会(FGDC)制定的美国空间数据元数据标准和空间数据转换标准

FGDC 是美国政府机构的一个协调性组织，其主要目的是在全国范围内促进对地理数据的共同开发、使用、共享和传播。

空间数据的元数据标准定义了一套数字化地理元数据的内容，并建立了相应的概念和术语，根据该标准的定义，元数据可以从以下七个方面对空间数据进行描述：

（1．1）标识(Indentification)：包括数据名称，开发者，数据描述的区域，专题，现势性，对数据使用的限制等；

（1．2）数据质量(Data Quality)：数据质量的定义，数据精度，完整性，一致性，产生该数据的原始数据以及处理过程；

（1．3）空间数据组织(Spatial Data Organization)：数字编码的空间数据组织方式，空间实体的数目，除空间坐标外其他的属性；

（1．4）空间参照(Spatial Reference)：数据采用的地图投影，存储格式（矢量还是栅格），水平与垂直的地球参照系，从一种坐标系统转换到另一种坐标系统的方法；

（1．5）实体和属性信息(Entity & Attribute Information)：数据中包括的地理信息，信息的编码方式，编码的意义描述；

（1．6）分发(Distribution)：如何得到数据，数据的格式，存储介质，价格等等；

（1．7）元数据参考信息(Metadata Reference)：该数据何时完成，由谁完成等信息。

2）空间数据转换标准(SDTS-Spatial Data Transfer Standard)是目前美国许多政府部门和商业组织所采用的交换格式标准。SDTS 是一个分层的数据转换模型，定义了数据转换的概

念、逻辑和格式三个层次，同时采用元数据来辅助数据转换和评价。

概念层建立了地理要素及其特征的模型，可以是矢量数据也可以是栅格数据，提供了地理要素的标准实体和属性的定义。逻辑层将概念化的地理要素转换成为逻辑化的模型、记录、数据项和子项，它提供了各种空间数据类型和关系的基础内容。SDTS 的物理格式层定义了与标准相符合的文件格式，以进行空间数据的转换。

SDTS 使得任意两种空间数据可以相互转换，并保证最小的信息损失，对于 NSDI（国家空间数据基础设施）的实现起到了决定性的意义。

3）加拿大 Mercator GIS 标准

Mercator 是以墨卡托投影来命名的，它是由加拿大政府组织、建立的一个国家 GIS 数据标准，其主要目标是：建立空间地理信息的标准，建立空间数据的存储库以及有关软件的开发。Mercator 定义了一个叫做 OGDI(Open Geospatial Datastore Interface)的程序接口，支持在客户机/服务器模式下访问空间数据。

如果不用人动手，梭子就能够编制，拨子就能够弹奏里拉，工头就不需要仆人，主人也不用指使奴隶了。

亚里士多德

第十八章 地理信息系统和社会

导读：地理信息系统应用的社会化是其发展的必然趋势，在社会化的GIS应用阶段，GIS技术在公众生活中被广泛使用，形成了GIS产业，这需要有相应的政策法规来规范产业运行，并且需要教育为GIS应用提供专业人才。在GIS技术影响社会的同时，一些社会因素也影响着GIS的发展。

1．GIS的社会化

1．1 实验室的GIS到社会化的GIS

在地理信息系统发展的初期，只是一门被少数科研人员所掌握的技术，他们利用各种GIS工具软件，甚至自己编写程序，处理空间数据，实现专业模型，得到服务于其研究工作的结果；而大量的GIS学者主要进行算法和数据结构方面的研究。对于一般公众，地理信息系统被认为艰涩难懂的，他们也无法意识到GIS技术可能的应用领域以及对社会生活带来的影响，这时的地理信息系统称为实验室的GIS。

由于计算机计算技术的发展和应用的普及，特别是网络技术和数据库技术的成熟，许多机构开始应用地理信息系统进行空间信息的管理，进而实现空间决策支持，以提高工作效率，减少企业运行成本，这种GIS应用方式称为“企业化GIS”，企业化的GIS可以概括为以下8个特点（陈子坦）：

1）**分布式的网络计算机环境**：基于网络的分布式计算使得企业各个部门可以共享资源，包括数据、设备、软件等。

2）**面向功能**：为了满足各个部门对GIS的要求，企业化的GIS需要以提供各种功能服务为主，保证各个部门按照其要求来使用。

3）**连续无缝的空间数据库**：为了达到对真实世界更有效、合理的管理，需要空间数据在存储时按照真实世界的形式连续的存储，而不是分割成为不同的图幅。

4）**数据的版本管理**：企业GIS在日常使用过程中，数据的变更是经常的，需要实施版本控制以保证数据的有效性，并避免混乱。

5）**不同层次数据和系统的共存**：企业的任务是多个层次的，这要求同时使用不同的系统和数据库。

6）**与多个外部数据库相连。**

7）**开放式的系统环境**：开放式的环境使得一个企业化的GIS与其它领域能够紧密结合起来，并且能够满足对系统进行扩展的需求。

8）**全面的GIS系统管理的专业人员**：企业化GIS对技术和管理人员有更高的要求，需要企业不断地对人员进行技术和管理培训。

在企业化的GIS应用中，GIS学者主要研究领域包括，基于分布式计算平台上的GIS实现，空间数据的共享和互操作，GIS项目管理等。

在企业化GIS应用的初期，其应用主要集中于一些经常要处理和使用空间数据的部门，包括土地，市政设施管理等等，这些部门的业务数据具有“强”的空间分布特性，如地块数据，管网数据等，采用GIS对其业务进行管理是“自然”的需求。随着信息技术应用的普及，许多其他领域的机构开始采用GIS,例如:商场可以采用GIS来分析不同区域客户的购物倾向;无线通信部门可以利用GIS来确定其发射台站的覆盖范围和覆盖效率；而交通部门可以使用GIS来分析交通流量。在这些机构中，空间数据只是信息系统中数据的一部分，或者信息系统的数据，如客户、雇员数据，具有“弱”*的空间分布特性，GIS与其它系统集成在一起，服务于其业务。

在GIS应用领域拓展的同时，对空间数据的应用层次也在提高，从简单的数据管理到复杂的空间决策支持，GIS可以满足各种机构不同的功能要求。遥感、GPS技术发展，为GIS提供了新的数据来源，多源的、海量的空间数据生成并被应用。上述的变化要求实现数据的共享、融合和分布管理，基于网络的分布计算平台为之提供了技术上的基础，同时GIS标准（包括ISO/TC211，OpenGIS等）的建立，通过规范空间数据模型、空间服务模型，确立标准的分布计算接口，保证了数据共享、融合和分布管理的实现。在上述的应用发展背景和技术发展背景下，GIS进入到社会化的应用阶段（图18-1）。

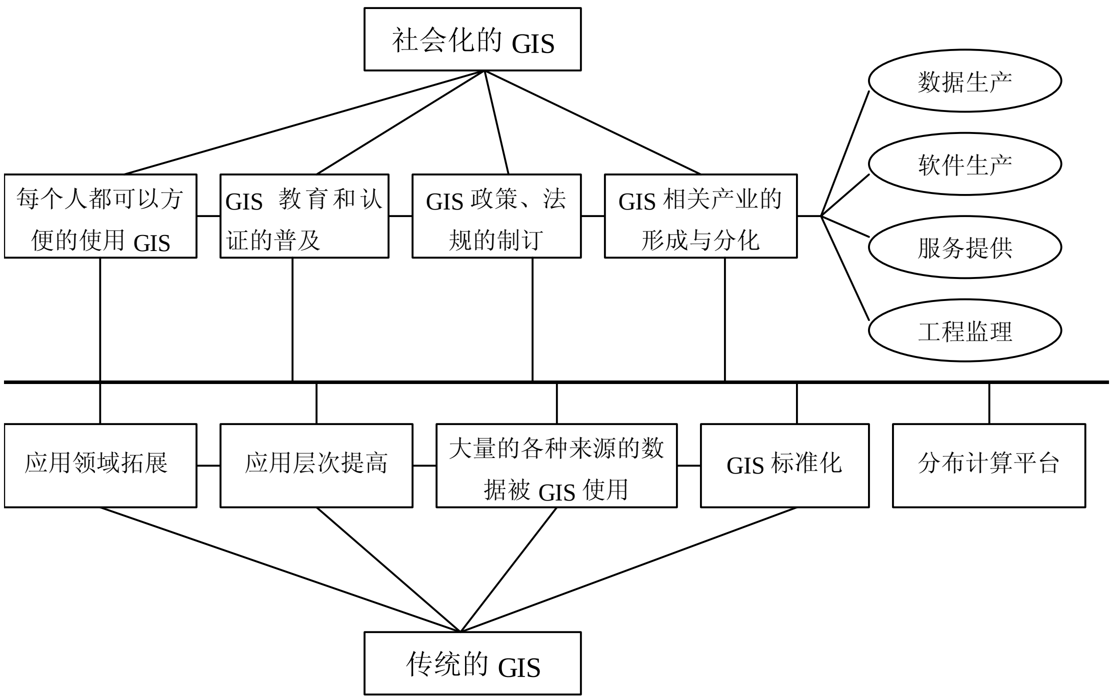


图18-1：地理信息系统的社会化

1．2 GIS 社会化的影响

GIS的社会化有一些其它的同义语，如“全球化”、“大众化”等等，它们实际上描述了GIS社会化的不同侧面。

“大众化”是GIS社会化的主要方面，它是指GIS技术已经融入到人们的日常生活中，迁

---

* “强”的空间分布特性是指数据的空间属性是所有属性中最重要的，也往往是数据使用者最关注的，如地块数据；反之，则称为“弱”的空间分布特征，如雇员数据，对于企业管理而言，雇员的居住地点属性显然不如工作技能，工资等属性重要。

移默化的改变着生活方式：

- 将地图存储在计算机中，使以往利用地图提供的定位、定向、导航功能可以通过GIS实现，而3S集成技术，可以使得定位导航功能更加自动化和精确；
- 利用GIS可以将信息按照其空间坐标组织起来，进行查询检索；进而
- 可以分析其空间分布特点，进行决策支持。

汽车导航、野外探险和旅游、银行信用卡管理、商家经营分析、保险赔偿分析等等，人们生活的各个方面，GIS都可以在其中发挥作用。当这种层次的应用从一个区域、一个国家扩展到全世界，全球的空间信息都基于GIS技术进行管理时，就实现了“数字地球”，可以认为，“数字地球”是GIS应用的极致，也是GIS社会化的顶点。

公众在日常生活中使用着地理信息系统，可是他（她）不需要了解任何GIS的知识，换言之，他（她）根本感觉不到GIS作为独立的技术在其中所起到的作用，正如一个基于数据库建立的自动取款机系统，当人们在存款、取款时，不需要了解数据库的查询、修改、提交、回送等具体功能的概念和实现，但确实，他用到了这些功能。这时，GIS应用和其他应用紧密结合在一起，已经成为人们日常生产和生活不可分离的一部分，正如一些学者所预测的，“GIS发展的将来就是没有GIS”。

从应用角度来看，GIS的社会化意味着每个人都可以方便的使用GIS功能。而从应用开发角度来看，GIS社会化的标志是GIS产业的形成与分化，形成专门的数据生产厂商，GIS平台/构件开发商，GIS集成商，GIS服务提供商以及GIS工程监理等等；相关GIS技术标准的确立，对于GIS产业的发展提供了基础。地理信息系统产业的发展，需要大量的专业人才，推动了GIS教育和认证的发展；此外需要政府和立法部门制订相应的政策法规，以保证产业运行的有序性（图18-1）。

2．GIS 社会化的其它问题

2．1 产业

随着GIS的社会化，地理信息已经发展成为一个巨大的产业，目前在美国，每年GIS应用项目多达1万个，创造产值在10亿美元以上。在产业发展和壮大的同时，也在发生分化，产生与地理信息相关的新的产业部门，产业的分化有利于形成规模经济，在整个社会范围内合理分配人力和财力资源，具体包括：

2．1．1 数据生产

在GIS应用中，地理数据的生命周期较硬件和软件都要长，并且地理数据的生产需要大量的时间，人力和资金的投入，有效生产和维护地理数据是地理信息系统社会化的前提。地理数据库的建立方式，除了由国家组织建设基础空间数据外，由专门的数据生产厂商录入和维护地理数据，并将之作为商品出售，可以提高数据的生产效率，避免数据的重复录入造成的浪费，降低GIS应用开发的成本。作为商品的数据的发售形式可以是光盘，磁带等，在因特网迅速发展的今天，通过网络发布数据的方式被越来越多地采用。

2．1．2 GIS 软件生产和系统集成

目前，有许多的公司在从事GIS软件的开发和系统集成业务，随着软件技术的发展，特别是软件开发过程中越来越多的采用构件技术，将形成GIS软件构件生产，GIS软件集成，GIS系统集成等多个部门（图18-2）。

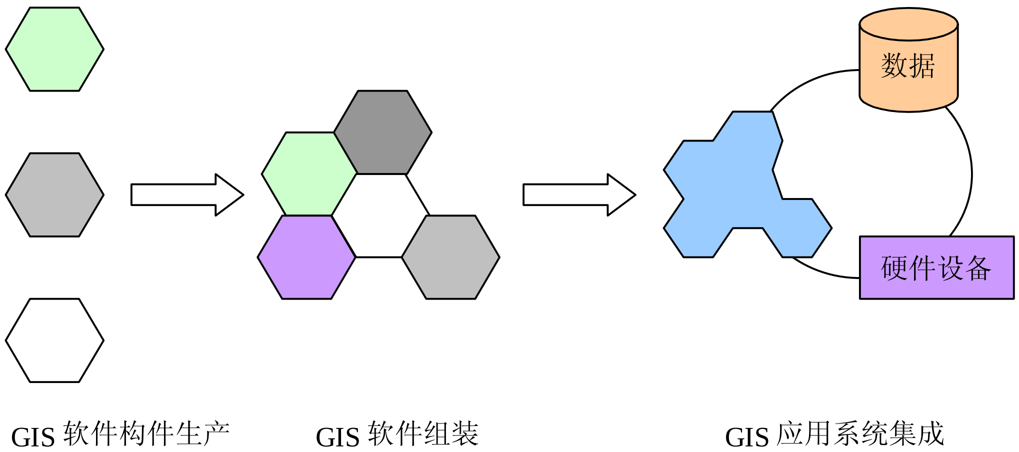


图18-2：GIS软件生产和系统集成过程中三个阶段的活动

GIS软件构件生产是按照公共的构件接口规范，生产GIS软件构件；而GIS软件组装则根据具体的应用需求，将构件组装成为应用软件，软件的具体形态可以是多种多样的；而应用系统集成则将组装完毕的软件与数据以及硬件设备集成，形成完整的GIS应用。

在传统的GIS应用阶段，GIS软件厂商提供了从软件生产到系统集成的各项服务，甚至包括数据的录入和维护；而在社会化的GIS应用中，这三个阶段活动分别由不同厂商提供，这样可以提高软件生产的效率，保证软件质量，最重要的，组装生成的软件可以方便地与其他应用系统集成，满足用户的各种需求。

2．1．3 GIS 服务

一个组织机构要使用GIS功能，除了通过开发或者购买建立自己的应用系统之外，还可以购买GIS服务，这样就需要有专门的GIS服务提供者。GIS服务包括的范围很广，最简单的如对客户提交数据的处理和信息提取，此外，GIS技术咨询，GIS认证等都可以划归到GIS服务的范畴，总之，提供GIS服务的目的就是为了帮助解决用户在建立和使用GIS的过程中遇到的问题，减少不必要的损失。

2．1．4 GIS 工程监理

许多机构在建立GIS应用，是通过项目招标，寻求开发商，然后双方签定合同，规定项目的具体内容，进而依据合同，开始进行系统开发。在项目开发过程中，开发方所依据的是用户的需求，但是由于用户方可能对具体的技术不理解，从而无法确定项目开发是否在按照自己的需求进行，因而也就无法对项目的进度、质量进行监控，如果到项目快要结束时才发现问题，会严重影响项目进度，甚至使项目失败。为了解决上述问题，一般规模较大的项目，都需要GIS工程监理，监理方由GIS领域方面的专家组成，他们对用户负责并了解用户需求，对项目的开发进行检查，可以及时发现问题，避免损失。

上面描述了在社会化的GIS应用阶段，所需要的GIS产业部门，随着应用的扩展和应用层次的加深，也完全可能出现新的部门。

2．2 政策

为了促进GIS应用的社会化，可以制订一些政策，提高GIS应用的广度和深度，此外，GIS社会化过程中，也需要一系列的政策法规，规范GIS应用以及产业的运作。

2．2．1 GIS 标准和规范

GIS标准和规范的制订便于实现空间数据的共享和互操作，指导GIS应用的建立，从某种意义上来讲，一个好的GIS标准是产业化的基础。标准的制订，既可以由一些标准化机构，如ISO、OpenGIS、OMG等组织；也可以由政府主持制订，并作为一项政策，在全国范围内颁布实施。通常，前者是指导性的，而后者具有一定的强制性。

除了制订GIS标准之外，一些政府性的GIS相关机构（如中国的测绘部门等）也可以参加一些国际性标准化组织的标准制订工作，并将国际化的标准特化，以适应本国的具体情况。

2．2．2 基础空间数据的建立和共享

国家基础空间数据建立，是一项浩大的工程，无法由任何单位单独完成，必须由国家进行组织，多个地方单位参加，协作完成。由于许多GIS应用是非赢利性的，这意味着项目组织单位无法承担高昂的数据购买费用，而重新组织数据录入无疑又是对人力和时间的巨大浪费，这种情况在一定程度上限制了GIS应用的发展。国家组织建立基础空间数据库，免费或低价提供给一些非赢利组织使用，如市政管理部门、科研部门等等，可以降低建立GIS的成本，有利于GIS应用的推广。空间数据标准是建立基础空间数据库的前提，另一方面，还需要制订相应的政策保证数据的共享。

2．2．3 其它领域应用 GIS 的规定

社会化的GIS应用，意味着许多机构主动地采用并建立GIS应用，以提高管理效率，同时，一些政策的制订，也有助于GIS技术被广泛接受，例如，可以规定一些城市规划的制订，自然灾害保险的赔付，必须要有GIS分析的结果作为依据，这样，一方面使得决策更加科学，另外可以加速GIS应用的发展。

2．2．4 GIS 产业的运行规范

随着GIS产业的发展和壮大，自然会出现与其他信息技术产业类似的问题，如不正当竞争等，需要有相应的政策进行规范。规范的制订和实施可以参照IT产业其他部门，同时，由于GIS在技术上特殊性，也往往需要GIS专家的参与。

2．3 法律

计算机技术的迅速发展和广泛应用，为人类社会提供了一种全新的生活方式，同时也带来了各种各样的，新的法律方面的问题，例如：

1）计算机犯罪

计算机犯罪指针对计算机系统的各种犯罪活动，如非法侵入计算机系统，盗窃或者非法使用计算机系统的数据，制造病毒等等。这些活动，既包括蓄意的或者出于好奇的破坏性活动，也包括利用计算机系统谋取非法利益的活动。

2）软件版权

随着信息技术产业的迅速发展，许多国家都认识到了计算机版权保护的重要性，并制定了相应的版权保护法规，以保护软件作者的知识产权，并保证信息技术产业运行的有序性。在信息技术领域，受到版权保护的对象包括程序、数据、文档、集成电路等等，不同的对象，其具体保护内容和侵权认定也有所不同，并且由于技术的不断发展，要求版权法也要进行相应的调整。软件（包括程序，数据和文档）由于其便于复制的特点，是版权保护的重点。

3）隐私权

计算机技术的发展和广泛运用对隐私权造成了较大的冲击，主要原因是通过计算机系统提供的便利，可以方便地收集和使用个人资料。特别是因特网的发展，该问题变得更为突出，因为用户在访问网站时，注册所填写的个人信息很有可能被非法使用。

4）网络法律问题

在计算机网络出现，特别因特网将全球网络连接成为一个整体以后，上述各个方面又发生了新的变化，出现了新的问题，如：网络作品版权，BBS 的言论自由，WWW 的内容控制，网络广告，网络病毒，网络上的不正当竞争，网络交易的安全性以及隐私权等等，这都需要对原有法律进行修改，以适应这些变化。

作为信息系统中的重要成分，地理信息系统在应用过程中，同样会遇到类似的问题，这些问题一般都可以适用普通的计算机法律条款，如 GIS 软件的版权等等。下面主要就地理信息系统中使用的空间数据版权进行讨论，这是因为空间数据在 GIS 应用中占有非常重要的位置，并且与通常的信息系统数据相比，空间数据有一定的特殊性。

美国是最早提出数据版权保护国家，对于数据版权的保护，其版权法做出了如下的规定：

1）将他人的版权作品纳入一项编辑作品内，必须取得这些作品权利人的许可，否则将构成侵权；

2）在纳入一项编辑作品之后，并不改变作品中各个组成部分原有的版权归属，编辑者只是拥有该数据的整体版权。

3）收集并编辑一些事实信息或者本身无版权的资料形成的数据，是否成为编辑者的原创作品，并由编辑者拥有版权，一般有两种观点，1）“辛勤收集(Industrious Collection)”原则，即尽管信息是已经存在的，只要编辑者在收集数据，并进行编辑处理，成为计算机可以检索形式的过程中，付出了经费、时间，使用了一定的技术手段，那么编辑者对其拥有版权；2）强调数据选择和组织安排的创造性，认为只有在信息的选择和组织安排方面体现了创造性的的数据才具有原创性。

上述的规定和原则，在处理一些数据版权纠纷时被经常采用，但是在一些具体的问题如将数据库的一部分下载到自己的计算机中以及对有版权的数据重新组织等行为的侵权认定还有待进一步的澄清。

国际上关于知识产权的伯尔尼公约（Berne Convention）的“文字和艺术作品”的定义中，明确规定了地图、地形图、与地理有关的三维作品（Maps, Topography, Three Dimensional

Works Relative to Geography）均属于知识产权保护的范围。各国在对 GIS 中的空间数据进行保护时，通常都援引与地图版权保护的法律条文。与一般的数据相比，GIS 数据可以有多种获得途径，包括纸质地图数字化，遥感图像解译，测量仪器以及现有数据的模型运算等都可以生成新的空间数据，此外，在开发 GIS 应用时，也可以购买空间数据；数据的处理复杂程度有很大的差别，既有简单的坐标变换，又有复杂的专业模型运算；数据的发布形式多样，如硬拷贝、磁盘、光盘、磁带以及 Internet 上的电子发布；数据应用的目的各异，空间数据可以应用于教学、科研、商业、公共服务、管理等各个领域。在进行版权保护时，需要综合考虑上述因素。

1）数据的获取

（1．1）利用测量设备，包括 GPS、平板测图仪等，得到的空间数据，无疑是其制作者（也包括数据制作单位）原创作品，作者对之拥有完全的版权；

（1．2）纸质地图数据的数字化，需要区分两种情况，即原地图是版权作品，及原地图是无版权资料。对于前者，数字化时需要取得地图权利人的许可。地图的数字化是一项非常繁重的工作，在数字化过程中，需要对原数据进行离散化和抽样处理，并且不同录入人员采用不同的录入方法，得到数据的精度也不同，即数字化后的数据不能等同于原纸质地图。从这个角度讲，数字地图的作者应该数据拥有完全的版权。

（1．3）数据购买只是获得数据的使用权，并没有得到所有权，购买数据的单位或个人将数据再私下转让给其他单位或个人，就构成了对权利人版权的侵犯。这一原则和普通的软件购买类似，软件购买可以通过许可授权做出更为明确的限制，而对于数据购买而言，购得的数据能否安装在不同的计算机中；为一个项目购买的数据能否应用于另外一个项目；数据使用许可能否象普通软件一样进行转让，尚需要进一步的探讨。

（1．4）遥感数据往往是通过购买得到，因而购买单位只是拥有使用权，但是对遥感数据进行处理，并得到专题数据（如植被覆盖图，土地利用图等），该过程融入了处理者的思想，所以一般来说，数据处理者拥有完全版权。

2）数据处理

只有在对有版权的电子数据进行处理，而处理者并不拥有其版权的情况下，才会涉及版权问题。数据的处理需要处理者投入时间和经费，并且体现了处理者的思想，数据处理完成后，原始数据在新的数据中并不显式地表现出来，根据“辛勤收集”和“创造性”的原则，在一般情况下，可以认为处理者拥有完全的版权。

但是，数据处理过程的复杂程度是不一致的，对于一些简单的变换处理，如投影变换，文件格式转换，图像处理中亮度、对比度变换，通常这些变换是可逆的，即由处理结果可以再生成原始数据。因而，处理者不能拥有结果数据的版权，否则会造成对原始数据权利人的侵权。在实际操作中，在简单处理和复杂处理之间进行绝对的区分是困难的，往往需要 GIS 专家的参与以进行判断。

3）数据的发布形式

空间数据可以有多种发布形式，包括地图，计算机可读的各种介质，Internet 等等，具有版权的数据，无论其发布形式如何，其版权都应该受到保护，除非权利人做出特别的声明。

4）数据应用目的

空间数据的使用可以有不同的目的，包括教育，科研，公共服务，商业等，一般对于非赢利目的使用数据，其限制可以适当放宽。

在 GIS 日益社会化，并且空间数据收集、数字化和加工成为 GIS 产业中的重要组成部分的条件下，制订针对空间数据版权保护的法规，对于保护数据生产商的利益，促进 GIS 应用的发展有重要的意义。

当然，随着 GIS 的社会化，带来的法律问题并不仅仅局限于数据版权，以下的一些问

题，如 GIS 产业的不正当竞争，空间数据的保密等等，都需要法律专家和 GIS 专家共同努力，提出合理的解决方案。

2．4 教育和评估认证

随着地理信息系统技术的发展，社会对GIS人才的需求也越来越大，同时，对于每一个GIS职员，他（她）所需要的GIS知识的数量和复杂程度也在不断加大。因而，许多GIS从业者在寻求进行培训和教育的机会。

目前，进行GIS培训和教育的机构主要包括职业学校、社区学院、大学和学院、GIS软件开发商、使用GIS的单位等。

这些机构进行培训的主要方式包括[T.A.WIKLE]：

1）工作室培训和短期课程培训班

工作室培训和短期课程培训班是一种不太正式的培训方式，在GIS发展的初期，这种培训方式占了很大的比例。工作室和短训班不需要特别的先决条件，通常讲述的是GIS软件。在一些情况下，对所有参与者，或者通过考试者发一份认证材料。

其课程可以是一般性的，也可以是针对某个应用或者软件的，前者一般由大学举办，后者则通常由GIS软件开发商组织（如ESRI的ArcView 4.0认证），此外，一个GIS应用机构在开展新的项目时，需要对员工进行培训；他们在雇用员工时，也比较看重短训班认证，因为这减少了他们培训工作量。

2）远程教学

远程教学是一种新的教学方式，它对于那些因为工作或者其它原因不能参加预定课程的人员来说，是一个较好的选择。参加远程教学的学员需要学习一系列课程，以获得文凭，这些课程的讲授资料可以通过视频信号或者因特网发送给远端的学员。

3）大学和学院的GIS课程

在GIS发展初期，由于昂贵的经费（软件，硬件，人员），很少有学校能够提供GIS教学，近十几年来，由于微机的广泛使用以及软件厂商针对教学的商业软件打折销售，使得大学和学院可以提供大量的、正式的GIS课程。到1996年，美国有600多所大学和学校开办至少一门GIS课程，中国目前也有许多大学的地理系或测绘工程系成立了GIS专业，另外一些相关的学科，如农学、土木工程、城市和区域规划、森林以及景观结构等也已经开始讲授GIS课程。

4）GIS认证

GIS认证一般需要学生已经完成了正规的GIS课程学习，在一些情况下，甚至需要获得GIS学位。象正规的学位一样，认证需要一个严格的评估过程，并且要大学或学院的教员参与监督。在一些情况下，研究院和公司合作，也可以进行GIS认证活动。

5）GIS学位

目前，特别是在加拿大和欧洲，许多学校已经开始授予地理信息系统和地球测绘(Geomatics)学位，但是美国能够授予GIS学位的学校较少，有些学校是通过计算机和数学系授予学士学位。

现在，中国的一些综合性大学（北京大学，南京大学）的地理系以及一些测绘专业大学（武汉测绘科技大学）都能够授予GIS专业的学士、硕士以及博士学位。

北京大学城市与环境学系地理信息系统本科专业及课程介绍：

北京大学城市与环境学系地理信息系统本科专业成立于 1998 年，学制四年，授予理科或工科学位，学生主要学习地理信息系统以及相关学科的理论、技术，使其能够在具体工作中应用 GIS，掌握一定的 GIS 应用开发技能。学习的专业课程包括：

1. 地理信息系统原理

地理信息系统的基本概念、特征、作用与应用现状；地理信息系统的基本构成及空间数据模型；地理信息系统应用的一般方法与流程；流行的地理信息系统软件介绍；地理信息系统应用方法与实例。

2. 空间分析

介绍如何利建立空间数据库和图形数据库以及空间数据与图形数据的管理与组织，在此基础上，借助地理信息系统软件进行区域、环境以及社会经济等空间查询与分析。

3. 地理信息系统设计

深入介绍地理信息系统的空间数据结构、空间模型的建立与分析；空间算法；网络 GIS 以及 3S 一体化的基本原理与方法；GIS 软件的设计与开发。

4. 地理信息系统工程

介绍地理信息系统工程的概念；数据的采集方法、规范、质量控制及数据组织与管理方法；地理信息系统二次开发技术；系统设计与开发的技术方案与流程；地理信息系统工程的建设与人员管理；工程的经济与社会效益分析；

5. 遥感原理

遥感的物理基础；遥感平台；传感器简介；目视解译方法；在土地利用、森林、农业、地貌、水文等方面的应用。

6. 遥感数字图像处理

介绍遥感数据的基本格式；遥感数字图像的几何校正、配准、变换等处理方法；人机交互方式的遥感解译与遥感分类；遥感图像的空间分析。

7. 测量与地图学基础

测量学的基本原理；GPS 的基本原理与使用方法；GPS 数据格式以及接口简介；地图与地图学的发展；制图符号；地图投影；标准分幅地形图的介绍；专题地图的信息表达；制图综合；地图出版过程。

8. 数字地形模型

介绍地形模型的概念与原理；数字高程模型的建立；坡度、坡向、粗糙度、幅照度等的计算方法；三维地形模拟；三维空间分析；在环境资源管理、土地利用、军事、地质地貌、水文等领域的应用方法。

9. 计算机概论

计算机的基本原理；软件与硬件介绍；操作系统的使用；多媒体简介；网络与通信简介。

10. 程序设计语言

以 C 和 C++语言为例，介绍计算机程序设计的基本原理。

11. 数据结构

介绍数据结构的基本原理，一些基本数据结构（线性表，串，树，图）等介绍以及其相关算法。

12. 数据库概论

介绍数据库的基本理论；数据库模型；关系数据库设计；数据库的运算与函数；数据库的基本操作；数据库程序设计；SQL 语言；面向对象数据库，分布式数据库；远程数据访问以及网络环境下的数据库系统开发与使用；数据库与地理信息系统的关系。

GIS应用的发展，需要大量的GIS专门人才，也需要大量GIS专业组织以提供各种服务，除了教育之外，还需要相应的评估和认证体系来保障人才和服务的质量。

目前，GIS领域对专业人员的能力和品质的评定，是依赖于一种建立于组织以及个人之间联系的“荣誉体系”。GIS应用的发展，需要一种更为结构化和客观的评估机制。其中，由于可以通过歪曲结果等方式滥用GIS技术，对于非道德的行为最需要关注。通常，在GIS领域，进行评估和认证的主要途径包括：

1）认证（Certification）

认证侧重于通过考试或者其它等价的评估过程，使GIS专业人员展示其能力和对GIS知识的掌握程度，从而得到一个精确的、可度量的结果。认证活动往往是非官方的，并且需要领域相近的专业协会的监督，通过认证也往往需要特定的学习和工作经验。

2）许可证（Licensing）

许可证也是保证GIS人员能力的一个途径，它又可以称为注册登记，是一个强制的过程，通过该过程政府给予一些个人从事某些领域的许可。要获得许可，申请人必须通过一个测试，并且可能需要完成一个认证程序或者获得学位。许可证的目的是保障公众的权利，当个人被确认从事了非法的或者不道德的活动，其许可证可能被收回。在很多国家，土地测量员需要有许可证。

3）授权（Accreditation）

授权不是针对个人，而是对GIS教育机构的认证过程，包括课程，教员的数量和质量，设施等等，例如：为了保证质量，授权过程规定了最少的课程数目。授权通过专业组织进行，其目的是保证GIS教育的质量。

GIS教育、评估和认证活动，使得地理信息系统应用组织能够得到足够的，合格的GIS专业人员，这些活动的进行同样需要相应的政策、法规的支持，其中评估和认证过程标准化是其中重要的方面，ISO/TC 211标准系列第22主题定义了这方面的标准。

3．社会对 GIS 发展的影响

GIS 技术既是一个工具，可以解决人们在生产生活中遇到的各种问题；同时，它也是一个社会过程，即 GIS 技术在影响和改变着社会的同时，社会也对 GIS 技术的发展施加着影响。

任何技术的发展都是一个演化的过程，该过程是多种可能的发展路径中的一个（图 18-3），在发展过程中，会遇到选择发展方向的各种状况，需要确定一个方向，而放弃另一个。通常影响决定因素包括：当时的实践、知识和社会条件等等。

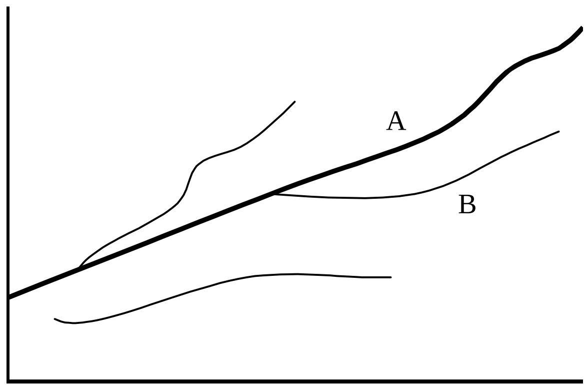


图 18-3：GIS 技术演化的图示[E. Sheppard]

下面两个条件对地理信息系统技术的演化起到了最重要的作用[E. Sheppard]：

1）首先是数字计算技术，GIS 的发展与计算机技术发展有着密切的联系，这意味着计算机技术的演化路径同样影响着 GIS 的发展，例如：基于图灵机逻辑结构的数字计算的采用以及高性能、小尺寸、低成本的微机的出现，对 GIS 的发展有着至关重要的影响。

2）其次，二战以后发展的一个重要的主题就是，各个机构能够有效的组织和管理生产以及有效的分发货物和服务。而计算机技术，包括 GIS 技术，是实现上述目标的主要途径，而具体的需求目标，也从另一个角度影响着计算机技术。

除了上述的一般的影响因素，具体影响 GIS 发展的因素包括：

1）社会，特别是军事部门，对于全球定位的需求的紧迫程度；

2）私营企业主导 GIS 开发的程度；

3）潜在的客户希望采用 GIS 解决的问题和他们预期的付出；

4）影响数据可用性和成本的因素；

5）地理学不再视为理论性的学科，这样可以利用地理学对 GIS 的开发和应用进行评估。

当然，投资于 GIS 技术最根本的目的是其经济价值，如果 GIS 开发主要是由私营企业进行的，这意味着 GIS 的发展是商业导向的，无论是开发数据产品还是平台系统，都需要在商业上获利。

上述的因素，决定着地理信息系统的发展途径，形成了现在 GIS 软件以及应用的样式。

一旦选择了一条技术方案，就具有一定的惯性，对以后的 GIS 产生巨大的影响。例如，基于层次结构的数据模型最早实现于哈佛图形实验室的软件 ODYSSEY，后来被 ESRI 的 ARC/INFO 所采用，成为 GIS 数据模型的主流技术方案，尽管面向对象的 GIS 很早就出现，并且在技术上更为先进，但是并没有在市场上取得成功，也就是说 GIS 的发展并没有经历这条途径。

球形的地代替了平面的地，引起了大地观念的依次彻底变化。这时人们无须扩展大地的圆盘以远远超过有人烟的地区，而认为有人烟的地区只包括地球的一小部分，更大的空余地面则可留待假说玄想去填充… …

阿尔夫雷德.赫特纳

第十九章 地球信息科学和数字地球

导读：本章介绍了 GIS 发展的一些最新的概念，包括地球信息科学，数字地球等等，这些概念的具体含义至今仍在变化。
数字地球与其说是一门技术，不如说是一个政策，它是 GIS 应用发展的顶点。最后介绍了国家空间数据基础设施，它是一个国家推广 GIS 应用重要的第一步。

1. 地球信息科学

1．1 几个相关概念

近十几年来，随着遥感，全球定位系统，地理信息系统以及计算机网络技术的发展，出现了一系列新的、意义相近的、与地理信息系统相关的名词，如地理信息科学（Geographical Information Science），地球测量（Geomatics，地球信息学[宫鹏]，地球空间信息学[李德仁]），地球信息学（Geo-Informatics），地球信息科学（Geo-information Science）等等，这些概念提出的时间还都不长，其含义存在交叉，目前国内对其确切的译名有些也存在着争论，下面介绍地理信息科学，地球测量的概念以及地球信息科学的概念和内容。

1．1．1 地理信息科学

地理信息科学是 1992 年 Goodchild 提出的，与地理信息系统相比，它更加侧重于将地理信息视作为一门科学，而不仅仅是一个技术实现，主要研究在应用计算机技术对地理信息进行处理、存储、提取以及管理和分析过程中提出的一系列基本问题，包括：

1）分布式计算
2）地理信息的认知
3）地理信息的互操作
4）比例尺
5）空间信息基础设施的未来
6）地理数据的不确定性和基于 GIS 的分析
7）GIS 和社会
9）地理信息系统在环境中的空间分析
10）空间数据的获取和集成等等

地理信息科学在对于地理信息技术研究的同时，还指出了支撑地理信息技术发展的基础理论研究的重要性。

1．1．2Geomatics

Geomatics 被定义为利用各种手段，通过一切途径获取和管理在空间基础信息生产管理过程中的空间参考数据部分的科学与技术[Gagnon]。Geomatics 还有其它的一些定义，这些定义有不同的表述，但是其要旨基本上是一致的：

1）定义空间参考基础

2）建立和使用对空间参照物体和现象进行定位和量测的方法、技术和工具

3）整合不同参考系统中的数据

4）提供合格数据

5）运用计算机技术改善数据的处理、存储和发行

Geomatics 包括大地测量、地籍测量、摄影测量与测深等传统测绘领域以及遥感和空间信息系统等新领域，可以认为是测绘学应现代社会对空间信息有极大需求这一特点提出的一个更全面、更综合的学科体系。

1．1．3 地球信息科学

随着以地理信息系统技术为核心的遥感、全球定位系统等技术的发展以及其间的相互渗透，逐渐形成了 3S 集成化技术系统，为解决区域范围更广，复杂性更高的现代地学问题提供了新的分析方法和技术保证。七十年代以来，由于整个人类社会面临的人口、资源、环境和发展等各方面的问题，逐渐开始重视全球变化（Global Change）以及可持续发展（Sustanable Development）等方面的研究，这两个方面的推动，最终促成了地球信息科学的产生。

地球信息科学以信息流的手段研究地球系统内部的物质流、能量流和人流的运动状态和方式（图 19-1），它由三部分组成，“地球信息学”是其理论研究的主体，“地球信息技术”是其研究手段，“全球变化和区域可持续发展”是其主要应用领域。[陈述彭]

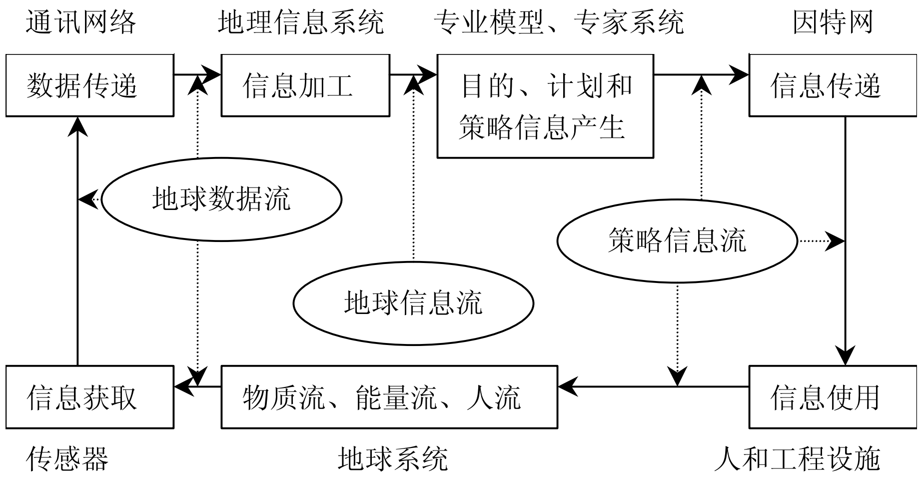


图 19-1：信息流和地球系统内部物质流、能量流和人流的关系

1．2 地球信息科学的研究内容

地球信息科学的本质是从信息流的角度来揭示地球系统的发生、发展及其演化规律，从而实现资源、环境和社会的宏观调控，作为其理论核心的地球信息机理研究包括：

1）地球信息的结构、性质、分类和表达；

2）地球圈层间信息传输机制、物理过程及其增益和衰减以及信息流的形成机理；

3）地球信息的空间认识及其不确定性与可预见性；

4）地球信息模拟物质流、能量流和人流相互作用关系的时空转换特征；

5）地球信息的获取和处理的应用基础理论等。

作为地球信息科学的研究手段，地球信息技术包括：

1）地球数据获取技术

用以从外部世界获得记录地球性质和状态的地球数据。遥感、全球定位系统等地球数据获取技术的发展已经形成了覆盖全球的监测运行系统，建立起多层次、立体的对地观测系统（EOS,Earth Observation System），是快速获取和更新地球数据的主要技术手段。

2）地球信息模拟技术

用于将地球数据转化为地球信息，包括地球空间信息的加工、再生以及获得有关某一地球空间问题解决的目的、计划和策略信息。

3）地球信息传播技术

用以实施空间数据的传递和信息的传播。

2．数字地球的产生背景以及其概念

数字地球（The Digital Earth,DE）的概念最早出现于1997年下半年，1998年戈尔在一次演讲中将数字地球正式提了出来。数字地球的提出是建立在人类科学技术发展成果的基础上，有其深刻的社会、经济背景。

数字地球的科学背景，包括地理信息科学，地球信息科学等一系列学科，这些学科的发展，为数字地球的产生打下了理论基础；对地观测系统以及计算机网络的发展为数字地球提供了技术支持；全球变化和可持续发展的研究对数字地球的出现起到了推动作用。（图19-2）

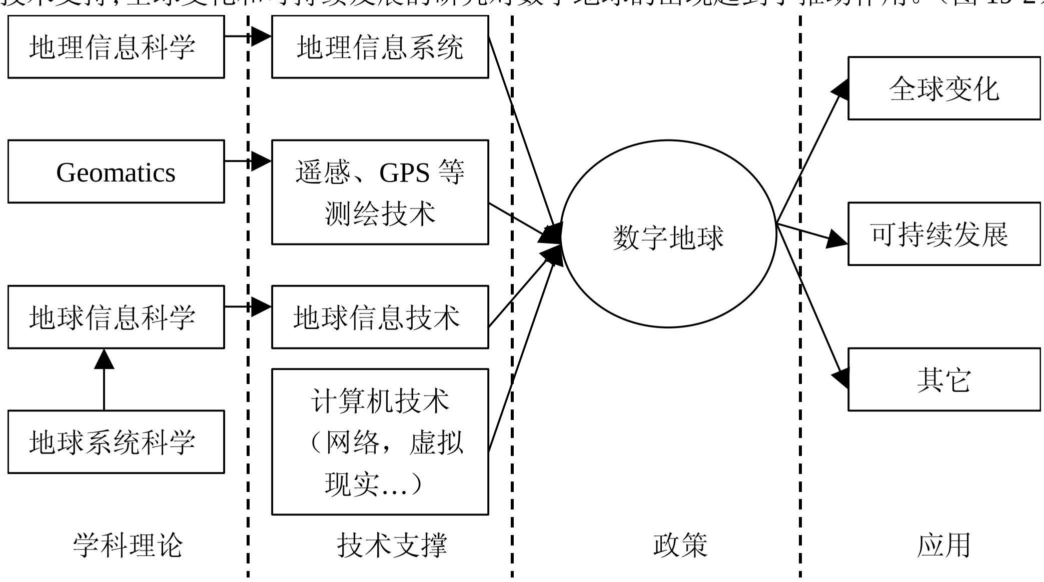


图19-2：数字地球和其它理论以及技术的关系

数字地球的基本概念，可以归纳为以下三个方面：

1）数字地球是指数字化的三维显示的虚拟地球，或指信息化的地球，包括数字化、网络化、智能化和可视化的地球技术系统；

2）实施数字地球计划，需要有政府、企业和学术界的共同协力参加。实施数字地球计划是社会的行为，需要全社会的关心和支持；

3）数字地球是一次新的技术革命，将改变人类的生产和生活方式，进一步促进科学技术的发展和推动社会经济的进步。

与地球信息科学、地理信息科学、Geomatics 等一系列学科概念以及遥感、GPS、因特网等技术名词相比，数字地球更象是一项技术政策，在一定程度上涵盖了上述的领域，“更加通俗易懂，是一个面向社会的号召，实质地说，数字地球就是要求地球上的信息全部实现数字化”[陈述彭]。

数字地球的基础技术，由遥感、遥测、地理信息系统、因特网等组成，遥感、遥测等提供获取数据的手段，地理信息系统进行数据的管理、存储、分析和表现，通过因特网实现数据的传输发布，这样构成了数字地球技术系统的基本框架（图 19-3）。

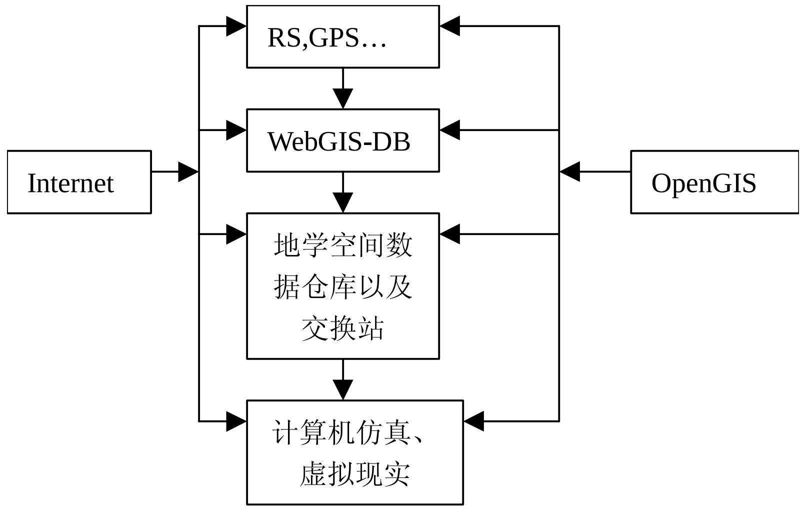


图 3：数字地球技术的基本框架示意图

数字地球：展望 21 世纪我们这颗行星，1998 年 1 月 31 日发表于加利福尼亚科学中心（艾尔。戈尔）

技术创新的新浪潮使我们能够大量地获得、存储、处理和显示关于我们行星的各种环境和文化现象信息。如此大量的信息构成了“地理坐标系”，它涉及地球表面每一个特定的地方。

利用卫星对地面遥感形成“地球空间”（轨道）信息流，并将其中未经加工有用部分的数据转变成为可供使用的信息。今天，我们经常发现我们能够得到比我们知道要做些什么更多的信息。地球资源技术卫星能够帮助我们更好地了解全球的环境就是一个典型的例子。地球资源技术卫星能够每两星期对地表进行一次完整的成像，而且它已持续收集了 20 多年。尽管对这类信息有着巨大的需求，但大多数图像还不能被人们使用。因此，它们只能被存储在电子数据库中。我们过去曾有过这样的农业政策，当成千上万的人们正面临饥饿死亡的时候，在中西部地窖里却存储着大量正在霉变的粮食。我们现在一方面表现出对知识的那种贪得无厌的饥饿，另一方面是大量数据目前仍然根本未被使用过。

问题之一是用什么方法把信息显示出来。有人曾经说我们是否能试用计算机术语来描述人脑，然而，人脑看起来比特率比计算机低，但却具有极高的分辨率。例如，研究表明由于人脑比特率低，因此，很难在短时间记住七条以上的数据。但是， 如果这些信息被排列在一个相互可以被识别的模型内如人类的脸或银河星系，人脑则能够同时吸收成千上万条信息。

过去我们通常利用 Macintosh 和 Windows 操作系统提供的桌面图形来进行（实现）人机对话的，事实上，这些工具现在已经不适合新的挑战。我相信我们需要一个“数字地球”，一个多种分辩率，一个三维的表述方式，使我们能嵌入巨大数量的地理坐标系数据。例如，

设想一位年轻孩子去本地的博物馆参观数字地球展览。在戴上显示头盔后，她看见地球从太空中出现，使用数据手套放大电子图像；使用越来越高的分辨率，看见大陆，然后是乡村、城市，最后是私人住房、树木和其他天然和人工的物质。她采用类似“魔毯骑乘”通过地形的 3D 视图，眼前被发现的行星区域引起她探究的兴趣。当然，地形仅仅是她能与之相互作用的多种数据之一。利用声音识别能力系统，她能够请求关于陆地覆盖、植物和动物的物种分布、实时天气、道路、政治范围以及人口的信息。她也能使她和其它学生收集的作为全球环境的一部分信息具体化。这些信息能被天衣无缝的融入数字地图或者地形数据。通过使用数据手套单击一个连接按钮，她能够获得她所看见的事物的更详细的信息。例如，为全家去黄石国家公园渡假作准备，她计划徒步去看天然喷泉，她刚刚阅读到关于野牛和洛矶山羊的知识。事实上，在她离开故乡的博物馆之前她能够进行虚拟旅行。

她可以不受限制穿越空间，同样也能穿越时间。在进行虚拟地去巴黎参观卢浮宫的场景旅游之后，她通过对时间地反向移动，学习法国的历史、细读被数字地球表面覆盖的数字化地图、新闻短片、口述历史、报纸等主要的原始资料。她向自己的电子邮件信箱发送一些有用的信息，以便今后学习。时间序列，从这里延伸到遥远的过去，能以天、年、世纪、甚至是地质学的纪元，因为那些场合有利于她对恐龙的了解。

显然，没有一个政府、企业或学术界能单独从事这项工作。像全球性网络，它需要成千上万的私人、公司、大学研究者以及政府组织的共同努力。虽然，作为数字地球的一些数据是公共信息。但数字地球也可以成为公司销售大量商用图像和有偿信息服务的数字化市场。它也可以变为一个“没有墙壁的实验室”，科学家可以在这里通过研究寻求理解人类和环境之间复杂的、相互依存的关系。

创建数字地球需要的技术

虽然这些情节似乎有些像科幻学小说，但大多数建造数字化地球需要的技术和能力已在这里或正在开发。当然，数字地球的能力将追随时间不断地进化。我们将有能力在 2005 年对 2020 年的数字化地球进行一次原始的比较。下面一些技术是需要的：

- 计算科学

直到计算机的出现，实验和理论两种创造知识的方法长期以来被限制，实验科学家对所要进行研究的现象的观察太艰难，这些现象不是太小就是太大；不是太快就是太慢。它们存在要么不到 10 亿分之一秒要么超过 10 亿年。另一方面，纯粹的理论不能预报像雷雨或飞机上的空气流动那样复杂的自然现象。随着高速计算机作为新工具开始被使用，我们才能模拟那些不容易观察到的现象，并且同时能更好地理解已经观察到的数据。按这种方法，计算的科学允许我们突破实验和理论科学的局限性。模拟和仿真将给我们对我们正在收集关于我们的行星的数据以新的洞察力。

- 大规模存储

数字化地球将要求存储 10^{15}字节的信息。稍晚于今年，国家航天航空局的星球计划将每天产生大量的有关地球的信息。幸运地是，我们继续在这个领域中制造戏剧性的改进。

- 卫星图像

早于 1998 年，政府已经许可提供分辨率图像处理服务的商业卫星系统开始运行这种服务能使地图绘制达到足够的精确的水准，而在原先绘制地图的详细资料只能通过空中摄影来获得。这种技术最先由美国情报组织开发，具有难以置信的精确性。当一家公司使用这种技术，它象一架威力无比的摄象机，在它所摄的图片上可以清楚地看到从伦敦到巴黎象汽车前灯的宽度大小的每一个物体。

- 宽带网络

数字地球所需的数据将由成千上万不同的组织维持，并非是单一的数据库。那意味在数字化地球中正在使用的服务器将需要与高速的网络连接。由于 Internet 网上通讯量爆炸的驱

动，电讯营运者已经在试验每秒 10Gbps 的网络。10Gbps 网络技术是 Internet 的下一代的技术目标之一。不好的消息是当我们绝大多数人的家中具有这类宽带之前，为什么数字化地球建立在诸如儿童博物馆和科学博物馆那样的公共场所的接口是必须的。

- **互操作性**

Internet 网和全球网之所以能够取得成功，关键在于达成统一的协议。如 Internet 协议。数字地球也同样需要互操作性。这样，应用软件产生的地理信息就能被另一些人阅读。GIS 产业界正在通过 GIS 协会寻找解决这些问题的途径。

- **Metadata**

Metadata 是指“关于数据的数据”，了解信息的名称、位置、作者或来源、数据、数据形式、分辨率等有助于建立其他的地理坐标信息系统。联邦地理数据委员会正在与产业界、州和地方政府一起开发非官方的 Metadata 标准。

当然，需要更进一步的技术进步来充分认识数字地球的潜力，尤其是在这些领域如自动解译图像、对来自不同渠道的数据的整合以及能够从网上找到和连接关于地球上某个地点的信息的情报机构等。

潜在的应用

全球地理信息系统的应用有时会被我们的想象力所限制。我们可以通过目前使用的 GIS 和数据传感器来感受一下这些应用：

- **虚拟外交**

为了支持波斯尼亚的和平谈判；五角大楼开发出一种虚拟现实场景，它能使谈判代表作一次虚拟的边界旅游。在谈判中，当塞尔维亚总统通过虚拟边界旅游看到原来的那条狭窄走廊是不切实际的之后，同意在萨拉热窝和伊斯兰教戈拉日德被占领土之间提供一条宽阔的走廊。

- **打击罪犯**

加利福尼亚莎利纳市的警方已通过使用 GIS 跟踪犯罪模型和团伙活动来减少青少年持枪暴力犯罪。通过收集关于犯罪活动的频率和分布的信息，使城市能够迅速重新部署警力。

- **保护生态多样性**

在 Camp Pendelton 的规划机构称：加利福尼亚地区预测人口将由 1990 年的 110 万增加至 2010 年的 160 万。这个地区被联邦或州政府立为被危害、威胁、或珍稀的而受保护的植物和动物超过 200 种。通过收集关于地形、土壤类型、年度的降雨量、植物、陆地使用以及所有权的信息，科学家模拟不同的地区增长计划对生态多样性产生影响。

- **预报气候变化**

模拟气候变化最显著的未知因素之一是全球森林的递减率。新罕布什尔大学的研究者与巴西的同事一起工作，通过卫星图像分析，能够监控亚马孙河流域的地形变化、这样他们就能确定该流域森林递减率。这项技术现在正在被世界上其它森林地区采用。

- **增加农业的生产力**

农民已经开始利用卫星图像和全球定位系统及时预防病虫害，并且把杀虫剂、化肥和水用到最需要它门的那里去。这被称为精耕细作，或称“按英寸渐进的耕作”。

- **未来方向**

我们有一个空前的机遇，来把有关我们社会和地球的大量原始数据转变为可理解的信息。这些数据除了高分辨率的卫星图像、数字化地图，也包括经济、社会和人口方面的信息。如果我们做得成功，将带来广阔的社会和商业效益，特别是在教育、可持续发展的决策支持、土地利用规划、农业以及危机管理等方面。数字地球计划将给予我们机会去对付人为的或是自然界的种种灾害，——或者说能帮助我们在人类面临的长期的环境挑战面前通力合作。

数字地球提供一种机制，引导用户寻找地理信息，也可供生产者出版它。它的整个结构

包括以下几个方面，一个供浏览的用户界面，一个不同分辨率的三维地球；一个可以迅速充实的联网的地理数据库以及多种可以融合并显示多源数据的机制。

把数字地球同万维网作一下比较是有建设性意义的(事实上它可能依据万维网和因特网的几个关键标准来建立)。数字地球也会象万维网一样，随着技术的进步以及可提供的信息的增加而不断改进。它不是由一个单独的机构来掌握，而是由公共信息查询、商业产品和成千上万不同机构提供的服务组成。就象万维网的关键是互操作一样，对于数字地球，至关重要的能力是找出并显示不同格式下的各种数据。

我相信数字地球的开发，首先需要建立一个由政府、企业和学术界共同参与的"试验田"。这块"试验田"将成为一些应用诸如教育、环境、解决与互操作性相关的技术问题、协作能力的联合以及解决一些政策问题如隐私权问题等的汇聚之地。当"试验田"获得成功并推广时，在更广泛的范围内可以通过接入高速网络与数字地球进行相互作用。而进入 Internet 将会感受到更多的限制。

可以肯定，数字地球不是一夜之间产生的。在第一阶段，我们应该把重点放在整合我们已有的、来自不同渠道的数据。我们同样应该把引导儿童的博物馆和科学博物馆连接到高速网络上如新一代的 Internet 网，这样我们的孩子就能够探索我们的星球。其次，我们应该努力去开发一张一米分辨率的数字世界地图。最终，我们应该使有关我们行星与我们历史的大量信息集聚到我们手指尖。

在数月前，我曾计划鼓励政府、企业、学术界和非赢利组织的专家们，制定一种实现这种前景的策略。共同努力，就能有助于我们解决社会所面临的许多最紧迫的问题；鼓励我们的孩子更多地了解他们身边的世界，并促进几十亿美元的工业经济增长。

3. 数字地球核心技术综述

为了解决数字地球中的数字化、信息化以及应用问题，需要研究以下关键技术：科学计算、海量存储、宽带网、卫星数据获取、元数据、互操作等等[戈尔]。下面对这些相关关键技术进行简单的描述。

1） 高分辨率卫星遥感数据的快速获取技术

卫星遥感是数字地球获取数据的主要手段，包括不同高度、不同分辨率的陆地卫星系列、海洋卫星系列、气象卫星系列以及小卫星系列，其分辨率从 1 米到 4000 米。遥感数据的处理包括辐射纠正、几何纠正、增强、特征提取、自动分类、自动成图、数据压缩等等，高分辨率卫星每天都要产生大量的数据，对这些数据的自动的、快速的处理以实时、准确地提取信息是实现数字地球信息获取的关键。

2） 地球空间数据的存储和处理

为了能够将地球上的信息进行数字化，除了要存储和处理大量的遥感数据之外，还包括图形数据，属性数据等等。实现对这些数据的查询检索，需要海量数据存储管理以及快速处理技术。目前分布式数据存储是海量数据管理的趋势，可以避免集中式系统带来的管理困难以及网络拥塞；而通常采用超大型计算机或者并行计算以实现快速处理。

3） 超媒体空间信息系统

数字地球的主要任务之一是通过因特网实现信息的共享和发布，主要通过 WebGIS 技术实现。此外，数字地球应用中包含大量多媒体数据，也需要在因特网上发布，形成超媒体空间信息系统。大量的数据在网络上传输，造成网络拥塞，这需要高带宽网络解决该问题。

4） 地理信息的分布式计算

地理信息的特征是分布的，并且具有基础性、共享性和综合性，分布式计算可以使得地理信息应用于社会各个领域。遵循 OpenGIS 规范，基于 CORBA（或 COM）体系结构，实

现地理信息信息的分布计算，是其解决方案之一。地理信息的分布计算服务包括：

（4．1）地理信息的共享领域服务

为社会各个领域的应用提供地理信息服务，提供共享和集成的基础。

（4．2）空间查询服务

包括通过元数据实现信息检索，属性、几何、空间关系等各个方面的查询。

（4．3）空间分析服务

实现空间信息提取，地理特征分析，图像理解，图像开发等服务。

（4．4）空间制图服务

包括空间坐标转换服务，地理注记服务，图像处理服务，特征综合服务，创建影象地图服务等等，其目的是为了直观的表现空间信息。

（4．5）地理信息的特定任务服务

为地理信息领域特定用户提供有针对性的服务，主要为地理信息的获取、地理信息的建立和维护服务，称为地理数据的管理服务。包括地理信息生产，地理信息生产管理，地图符号管理，信息开发利用和分析，订购和跟踪，信息存储和检索系统，信息交付，信息再现和复制，用户支持。

5） 无比例尺数据库

无比例尺数据库是指以一个大比例尺数据库为基础数据源，在一定区域内空间对象的信息量随比例尺变化自动增减，即可以由大比例尺空间数据自动生成较小比例尺的数据。无比例尺数据库的技术关键是自动制图综合，在该领域目前已有许多研究，但是都难以达到令人满意效果。

6） 空间数据仓库

空间数据仓库是指支持管理和决策过程的、面向主题的、集成的和随时间变化的、持久的和具有空间坐标的地理数据的集合。数据仓库的主要任务是将来源、结构、格式不同的原始数据，首先对其进行标准化、过滤与匹配、精化、标明时间戳和确认数据质量的处理，即求精过程；然后再根据任务的需要，进行数据的集成与分割、概括与聚集、预测与推导、翻译与格式化、转换与再映象处理；最后进行数据仓库的建模、概括、聚集、调整与确认及建立结构化查询等等。空间数据仓库的目的，是为了处理积累的海量空间数据，抽取有用信息，并提供决策支持。其结构框架包括：数据源，Metadata 数据源，Metadata 互操作协议，数据抽取求精，Metadata 创建与浏览数据仓库，存取与检索，Metadata 管理及查询与分析等。

7） 空间数据融合(Fusion)

空间数据融合是指多种数据合成后，不再保存原来的数据，而产生了一种新的综合数据，如假彩色合成影象。数字地球的多种数据融合，包括多种分辨率数据，多维数据以及不同类型数据的融合，并且需要将融合得到的数据进行可视化表现，通常是将数据叠加在数字高程模型上，形成三维立体景观影象。实现数字地球中的空间数据融合，需要地理数据互操作以及高速网络的支持。

8） 虚拟现实(VR,Virtual Reality)技术

虚拟现实技术，是指运用计算机技术生成一个逼真的，具有视觉、听觉、触觉等效果的可交互、动态的“世界”，人们可以对该虚拟世界中的虚拟实体进行操纵和考察，用户与虚拟现实系统的交互利用数据手套、数据头盔、数据衣等进行，而 VR 系统通过视觉描绘器，听觉描绘器，触觉描绘器使用户产生身临其境的感觉。在数字地球中采用虚拟现实技术，可以非常真实地表达现实地理区域，而用户可以在所选择的地理带内和外自由移动。目前，GIS、虚拟现实以及 Web 技术相结合的方式之一是 VRML(Virtual Reality Modeling Language，虚拟现实造型语言)，通过用 VRML 描述 GIS 信息，可以在因特网上发布空间三维数据供用户浏览。

9）元数据（Metadata）

在创建数字地球的过程中，全球范围内对数字地理信息的需求越来越大，许多单位和个人开始生产、处理和修改地理数据；另外，在计算机信息系统中，在采用模型对地理实体进行研究时，为了保证信息不被误用，需要通过 Metadata 对数据进行详细的描述，这样不仅数据生产者能够充分描述数据集，用户也可以估计数据集对其应用目的的适用性。所以，随着地理空间数据生产者和使用者的增加，使用 Metadata 来描述数据，将成为一个必然的趋势。

综上所述，数字地球技术的关键是实现海量数据获取、存储管理、处理和信息提取、信息共享以及信息表达，下面给出了它们与上述具体技术的关系（表 19-1）。

表 19-1：数字地球技术概括

	海量数据获取	存储管理	处理和信息提取	信息共享	信息表达
高分辨率卫星遥感	_				
地球空间数据的存储和处理		_	_		
超媒体空间信息系统				_	_
地理信息的分布式计算		_	_	_	
无比例尺数据库		_	_		_
空间数据仓库		_	_		
空间数据融合			_		_
虚拟现实技术					_
元数据	_	_	_	_	

说明：“_”表示该具体技术支持相应的实现目的。

4．国家信息基础设施和国家空间数据基础设施

正如前面所描述的，数字地球作为一项技术政策，其建立必须要有政府的参与，除了组织和支持相关技术领域的研究之外，最重要的方面就是建设国家信息基础设施（NII,National Information Infrastructure）以及国家空间数据基础设施(NSDI,National Spatial Data Infrastructure)。

4．1 国家信息基础设施

国家信息基础设施是一个能够给用户随时提供大容量信息的，由通信网络、计算机、数据库以及日用电子产品组成的完备的网络系统。目前全球被广泛采用的信息基础设施就是因特网，而 Web 服务无疑是因特网上最重要的应用。国家信息基础设施将促进经济的发展，其预期效益如下：

1）推动新技术发展，如半导体、高速网络及其软件；

2）形成大型产业，加速经济发展；

3）促进电子商务的发展；

4）协助解决医疗保健问题，降低医疗费用；

5）促进科技研究；

6）促进教育事业；

7）为全国公民服务。

4．2 国家空间数据基础设施

国家空间数据基础设施(NSDI)是国家信息基础设施之后的又一个国家级信息基础设施，其目的是为了协调基础地理空间数据集的收集、管理、分发和共享的基础设施。空间数据基础设施主要由四个部分组成：数据交互网络体系、基础数据集、法规与标准、机构体系。从技术的角度来看，其内容主要有：空间数据标准，基础空间框架数据，空间数据交换网络以及元数据等等。

根据计算，目前全世界的数据中，有 80%的数据包括空间参考数据内容，也就是说，地理信息已经渗透到各个部门和学科。许多组织和单位都需要利用空间数据进行业务生产或科学研究，这样，确立数据标准，依据相应的制度和法规，指导空间数据的录入和管理，以实现数据共享，可以进一步推动地理信息的使用，使地理信息应用单位不需要重复录入空间数据，减少其工作成本，这是建设国家空间数据基础设施所带来的最大裨益。另一方面，数字地球的建设是高技术难度，需要大量资金投入的工程，而国家空间数据基础设施可以视为一个国家数字地球工程重要的第一步，因为它构造了空间数据库，确立了相应的政策、法规和标准，这也是数字地球的实现框架中所必需的。

目前，我国 NSDI 的建设包括空间信息的收集、管理、协调和分发的体系和机构；空间数据收集系统；地理空间数据集 Metadata 和空间信息交换网络；基础空间框架数据以及地理空间数据标准。

1）空间信息的收集、管理、协调和分发的体系和机构：制订国家空间信息设施的规划、政策、标准和法规，建立层次化的机构体系，协调各个部门（包括测绘、土地、环境、开发、农业等部门）的协作和权益，研究相关技术，进行项目管理等等。

2）空间数据收集系统：包括一系列遥感卫星或小卫星组成的遥感卫星体系。

3）地理空间数据集 Metadata 和空间信息交换网络：包括国家基础地理信息中心，中国科学院，国家信息中心，北京大学等许多单位已经开展了 Metadata 标准方面的研究，但是还需同国际上相关工作的接轨。空间信息交换网络主要利用了国家信息基础设施，包括中国公用分组交换数据网（CHINANET），中国互连网，金桥网（CHINAGBN），教育科研网等。

4）基础空间框架数据：国家测绘局进行了基础空间框架数据的输入和建库工作，其中包括全国 1:25 万地形数据库、全国 1:25 万地名数据库、全国 1:25 万数字高程模型、全国 1:100 万地形数据库、全国 1:100 万地名数据库、全国 1:100 万数字高程模型库、全国 1:400 万地形数据库、全国 1:400 万重力数据库、部分地区 1:5 万 4D 产品系列、全国七大江河 1:1 万 4D 产品系列等等。

5）地理空间数据标准：国家测绘局以及其它相关部门建立一系列空间信息标准，并已经发布实施，包括“地理格网”，“国土基础信息数据分类与代码”，“林业资源数据分类和代码”，“全国河流名称代码”等等。此外包括涉及空间数据交换格式，椭球体和投影等各个方面的标准也正在制订之中。

附录一 GIS 的计算机基础

导读：本章讲述了掌握 GIS 功能所需要的计算机基础知识，包括：
1）计算机组成原理：简单概述了计算机的发展历史，介绍了计算机的硬件组成，重点是各种输入/输出设备，这是 GIS 数据录入和制图输出所必需的，此外，虚拟现实设备——包括数字手套、头盔显示器的介绍将有助于理解数字地球部分的虚拟现实技术。
2）数据库知识：介绍了数据库，数据库管理系统的基本概念以及层次、网状、关系三种数据模型，并针对关系数据库讲述了其基本操作和 SQL。
3）数据结构和算法：介绍了数据结构的基本概念，以及对于 GIS 软件实现非常重要的两种数据结构——树和图，最后给出了算法的定义以及算法效率的衡量指标，可以作为 GIS 算法设计的指导。

1．计算机组成原理

1．1 计算机发展历史和展望

电子计算机的发展应该追溯到第二次世界大战期间，由于当时急需高速、精确的工具来解决弹道计算问题，1943 年 4 月，宾夕法尼亚大学摩尔电工学院开始试制一台被称为 ENIAC（Electronic Numerical Integrator and Computer）的电子数字计算机，并于 1946 年 2 月 15 日制造出了这台机器。1946 年 6 月诺伊曼（Neumann）等发表了一份《关于电子计算装置逻辑结构初探》的报告，这份报告，提出了以二进制和程序存储控制为核心的通用电子数字计算机体系结构原理，奠定了电子数字计算机体系结构的基础，从此翻开了计算机发展的新篇章。

在电子计算机问世后，它所采用的基本电子元器件已经经历了电子管——晶体管——集成电路——大规模集成电路四个发展阶段，通常称为计算机的四代（图 A1-1）。

1）电子管计算机时代，从 1946 年到 1958 年左右。这代计算机因采用电子管而体积大，耗电多，运算速度低，存贮容量小，可靠性差。虽然这个时期的计算机原始而笨重，但却确立了计算机发展的技术基础，例如，二进制、自动计算及程序设计等。

2）晶体管计算机时代，约为 1958～1964 年。这代计算机比第一代计算机的性能提高了数十倍。软件配置开始出现，一些高级程序设计语言相继问世，外围设备配置也由几种增加到十几种。除科学计算外，开始了数据处理和工业控制等方面的应用。

3）集成电路（IC,Integrated Circuit）计算机时代，约从 1964 年到 1970 年。这代计算机主要由中、小规模集成电路组成。这种电路器件是在一块几平方毫米的芯片上集成了几十个到几百个电子元件，使计算机的体积和耗电有了显著减小；而计算速度和存贮容量有较大提高，可靠性也大大提高，计算机软件配置进一步完善；有了操作系统，系统结构方面有了很大的改进。计算机机种多样化、系列化，并和通讯技术结合起来，使计算机应用进入许多科学技术领域。

4）大规模集成（LSI,Large Scale Integration）电路计算机时代，从 70 年代开始到现在。大规模集成电路是在一块几平方毫米的半导体芯片上集成上千个到十万个电子元件，使得计算机体积更小，耗电更少，运算速度提高到每秒几百万次。计算机可靠性也进一步提高。

七十年代初，出现了微处理器。它是把计算机的运算器、控制器制作在一片大规模集成电路芯片上，从而可以把处理器和半导体存贮器芯片以及外围接口电路芯片等组装在一起构成微型计算机。微型机体积小，价格便宜，灵活性大，使得计算机应用迅速发展，开始了个人用计算机的时代。

目前，计算机技术正在继续向巨型、微型、网络和人工智能等几个方向发展：

1）为满足尖端科学研究的需要，还必须发展高速、大存贮容量和强功能的巨型机；

2）计算机另一个发展方向是要研制价格低廉、使用灵活方便的微型机，以适应各种应用领域。

3）计算机网络是计算机的又一发展方向，计算机网络提高了计算机系统资源，特别是信息资源的综合利用，把分布在许多地区的计算机系统，特别是分布在各地的信息资源联结在一起，组成一个规模更大、功能更强、可靠性更高的信息综合处理系统。

4）美国、日本等国正在研制第五代“智能”计算机，它不是注重数学运算，而是注重于逻辑推理或模拟人的“智能”。

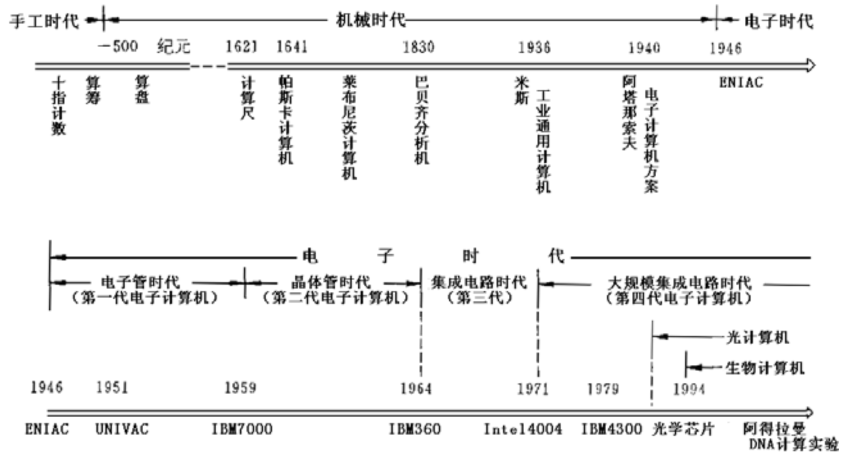


图 A1-1：计算工具的发展概况

1．2 计算机的基本组成

按照诺伊曼理论建立起来的当代计算机，应当具有输入/输出功能、存储记忆功能、计算功能、判断功能和自我管理功能。从功能模拟的角度，Neumann 计算机通常由与上述功能对应的功能部件组成，这些部件主要包括输入/输出设备、存储器和中央处理单元（CPU,Central Processing Unit）。它们之间的关系如图 A1-2 所示。

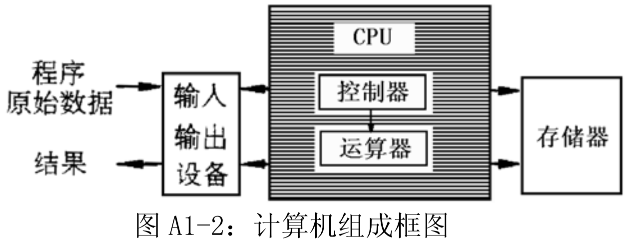


图 A1-2：计算机组成框图

1）输入/输出（I/O）设备

输入输出设备是接收外部信息（如输入原始数据和程序）或用来向外部输出信息（如计算结果）的功能部件，包括打印机、显示器、键盘、磁带机、扫描仪、鼠标器、光笔、触摸屏、条形码阅读器等。

2）中央处理单元（CPU,Central Processing Unit）

CPU是计算机的核心，它主要由运算部件、控制器、寄存器组所组成：

（2．1）控制器

控制器的主要功能是按时钟提供的统一节拍，把程序中每一条指令所含的各基本操作进行时序分配，并发出相应的控制信号，驱动各部件按照节拍有秩序地完成程序规定的操作内容。

（2．2）运算部件（ALU）

运算部件是直接进行数据变换与运算的部件。运算部件主要由逻辑电路构成的加法器组成。加法与逻辑运算是运算器最基本的操作，由它们可以进一步实现四则运算（加、减、乘、除）和逻辑操作（逻辑运算、条件运算等）。

（2．3）寄存器组

运算部件进行计算需要输入两个操作数，并产生两个输出：结果和运算特征。运算特征也称运算状态，如操作结果是否为零、是正还是负、有无进位、操作是加还是减等等，取得这些操作特征的目的是为了决定下一步的操作。

3）存储器

存储器是计算机的记忆装置，用以保存程序、原始数据以及中间结果。目前，计算机基本上采用线性地址存取方式。每一个地址对应一个存储单元，存储单元可以按位（bit）或按字节（8bit）、字（16bit）、半字（8bit）、双字（32bit）等编址。在按字节编址的情况下，每个存储单元存储一个字节（1Byte）信息。

存储器中存储单元的数量称为该存储器的容量，它是评价计算机功能的重要指标之一。存储器容量愈大，所能存储的信息就越多，可处理的问题的复杂度就越高。容量、价格、存取速度是评价计算机存储器的三大指标。但三者之间又互相制约：容量大，存取速度就要低；采用存取速度高的元件，成本就高；成本高，就不允许做得容量太大。为此，通常采用分级存储方式来解决这三者之间的矛盾。最基本的分级存储结构是两级存储，即把存储器分为主（内）存储器与辅助（外）存储器两级。主存储器采用半导体存储器，辅助存储器采用磁介质存储器。磁盘就是广泛使用的一种辅助存储器。为了进一步提高计算机的性能，在主存和CPU之间又增加一级比主存速度更高的高速缓冲存储器Cache，形成三级存储体系。

1．3 存储系统

早期的诺伊曼计算机是以运算器为中心的，系统内各部件间的信息传送都要经过运算器。随着计算机应用的深入和外部设备的发展，内存与外存等外部设备之间的信息交换日益频繁，为适应这一情况，形成以存储器为中心的系统结构，主存同外部设备之间的信息交换不再通过运算器。共享主存的多处理机的出现，更加强了存储器作为计算机系统的中心地位。这时，存储器除了要向一台或多台高速运行的CPU提供所需的指令和数据外，还要同并行工作的外存及其它外设和终端等设备交换信息。存储系统的特性，已经成为影响整个系统最大吞吐量的决定性因素。

广义地讲，在一定条件下，物质性质的改变，就是对过程条件的记忆，如果这些物理性质可检测并且与其相应过程条件之间有确定的一一对应关系，则可用做记忆元件。基于二进制逻辑的电子计算机所要求的记忆元件应当有两个明确定义的物理状态，以分别表示两个逻辑值，并且这两个状态可以被检测并转换成电信号。信息的存取速度取决于测量与改变元件的记忆状态所需的时间，能满足这一要求的物质有机械的、磁的、电子的、光学的、化学的和生物的等等。

1．3．1 主存储器组成

目前，主存储器中所使用的记忆元件是电子的，即半导体的，包括：

1）半导体 RAM（Random Access Memory）记忆元件

随机存取存储器 RAM 要求能随机地对存储器中的任何单元进行存取，且与存取的时间和该单元的物理位置无关。具体地说，它要求元件有如下记忆特性。

- 有两种稳定状态；
- 在外部信号的激励下，两种稳定状态能进行无限次相互转换；
- 在外部信号激励下，能读出两种稳定状态；
- 可靠地存储。

2）半导体 ROM（Read Only Memory）记忆元件

ROM 是一种在机器运行过程中只能读出、不能写入信息的无源存储器，是一种非易失性器件。它所存储的信息是用特殊方式写入的，主要用于存储器经常要用的一些固定信息。

3）闪速存储器（Flash memory）

1．3．2 辅助存储器

辅助存储器是主存储器的后援存储设备，用以存放当前暂时不用的程序或数据。对辅助存储器的基本要求是：容量大、成本低、可以脱机保存信息。目前主要有磁读写、光读写两类，如磁盘、磁带、光盘、光磁盘等。

1）软盘存储器

软盘存储器由软磁盘、软盘驱动器、软盘适配器三部分组成。它们是目前个人计算机中应用最为广泛的一种辅助存储器。

2）硬磁盘存储器

硬磁盘的盘片以铝合金为基体，因而“硬”，但存储原理与软磁盘相同，相对软盘，硬盘存储器存储容量更大，访问速度更快。

3）磁带存储器

磁带存储器是最早应用的磁表面存储器，其特点是存储容量大、价格便宜。

4）光盘存储器

由于多媒体的发展，这促使了光盘技术的迅速发展。光盘存储器是把激光束聚焦成 lum 左右的微小光点，使之能量高度集中，在记录介质上产生物理或化学变化而存储信息的。读出时，激光束在介质上扫描，根据反射光的变化判断记录的数据。

光盘存储技术记录密度高，存储容量大；可长期（60 年～100 年）保存信息；成本低廉，易于大量复制；存储密度高，体积小，能自由更换盘片；是很好的大容量存储技术。但是光盘的数据存取速率比磁盘低，目前一般为 50～150MB/S，因此还不能完全取代磁盘。

5）磁盘阵列 RAID

RAID（Redundant Arrays of Inexpensive Disk）是并行处理技术在磁盘系统中的应用。它把多台小型的磁盘存储器（或光盘存储器）按一定的条件组织成同步化的阵列，利用类似于存储器中的多体交叉技术，将数据展开存储在多台磁盘上，提高了数据传输的带宽，并用冗余技术提高了系统的可靠性。

1．4 输入/输出系统

输入/输出系统是计算机主机与外界交换信息时需要的硬、软件设备的总称，简称外设系统。一般说来，外设系统的硬件由以下几个方面组成，这里主要介绍外部设备。

1）外部设备：围绕主机而设置的各种信息媒体转换和传递的设备。

2）设备控制器与接口：控制主机与外部设备之间的信息格式、交换过程、外部设备运行状态的硬、软件，也称设备适配器，它与外部设备的特性有关。

3）I/O 总线；主机与外部设备之间的信息传送通路。

1．4．1 外部设备及其分类

“外部设备”也称为外围设备。它们是指计算机系统中，除主机以外，直接或间接与计算机交换信息、改变信息媒体或载体形式的装置。从使用的角度，外部设备大致可以分为如下三类。

1）人——机交互设备

人一机交互设备，就是用户和计算机间交流信息的设备，其功能是把用户可以识别的信息媒体，转换成计算机可以识别的信息，如键盘、图形扫描仪、摄像机、语言识别器等；或者把计算机处理的结果信息，转换为用户可以识别的信息媒体，如打印机、显示器、绘图仪、语音合成器等。

2）机——机通信设备

机——机通信设备就是一台计算机与其他计算机或别的系统之间通信的设备，如两台相同型号或不同型号之间的计算机利用电话线路进行通信时，所需的调制解调器（MODEM）以及用计算机进行实时控制时的数/模（D/A，Analog/Digital）——模/数（A/D）转换设备等。

3）计算机信息的驻在设备

计算机信息的驻在设备，即计算机的外存储设备，如磁盘、光盘、磁带等。这些已在前面介绍过了。本节主要偏重于介绍人——机交互设备。

1．4．2 字符输入/输出设备

1）键盘

字符输入设备的实质是将要输入的字符转换成相应的 0、1 码。目前，键盘是最重要的字符输入设备，键盘的基本组成元件是按键开关，它的种类很多，一般可分为触点式和无触点式两大类。

2）打印设备

打印设备是一种硬拷贝设备，它的作用是将输出信息打印在纸上，产生永久性记录。打印设备种类繁多，有多种分类方法。按印字原理分类，可以分为：

- 击打式:打印过程打印头要撞击纸。击打式打印机又分为活字式打印和点阵式打印;
- 非击打式：采用电、磁、光、喷墨等物理、化学方法印刷字符，打印过程纸不被撞击。如激光印字机（其技术来自复印机）、喷墨印字机等。

按工作方式可以分为：

- 串行打印机：逐字打印；
- 行式打印机：一次输出一行。

1．4．3 图形/图像设备

1）绘图仪

绘图仪与图形显示器相似而又不同，不同处在于它是输出永久性图形的设备，而图形显示器是输出过程图形的设备，并且它们的结构元件不同；相似之处在于，它们形成图形的原理相似，即按形成图形的元素，绘图仪也可以分为向量绘图仪和点阵绘图仪两类。

（1）向量绘图仪

向量绘图仪，又称笔绘图仪，构成向量图形的基本元素是直线段。

（2）点阵绘图仪

点阵绘图仪又称无笔绘图仪，组成点阵图形的基本元素是点，或称象素。属于点阵绘图机的有静电绘图仪、喷墨绘图仪、热敏绘图仪和激光绘图仪等。

2）摄象机和扫描仪

（1）摄象机

摄象机是最直接的图像输入设备，它能把所摄图像经数字量化后变成数字图像存人磁带、磁盘或光盘，以备放映。

（2）扫描仪

扫描仪是能够全面而快速输入数据、文字和图形以及图像的输入设备。扫描仪从结构原理上可分为两大类，一类是 CCD 作为光敏元件，另一类是以光导纤维作为光的传导元件。在地理信息系统数据录入中，通过扫描仪得到图像数据，然后进行跟踪矢量化，是快速获取数据的重要手段。

1．4．4 定位及拾取设备

定位/拾取设备通过指点来读取（位于屏幕或图表、图形上的）坐标，以画出或修改图形。按所拾取的坐标分为两类：拾取绝对坐标，如光笔和数字化仪；拾取相对坐标，如鼠标器、跟踪球、操纵杆等。

1）光笔

光笔是一种输入设备，用来检测信号，因为外形像支笔，所以叫光笔。其前端装有光敏器件，后端用导线接到计算机上。当光敏端的笔尖接触屏幕时，产生的光电信号向计算机发出中断脉冲信号，此瞬间显示存储器的地址就是光笔所指位置，计算机按操作人员的命令作出响应或画图、编辑和修改。

2）触摸屏

触模屏是一种能对物体的接触或靠近产生反应的定位设备。根据采用技术之不同，触摸屏分为五类：电阻式、电容式、表面超声波式、扫描红外线式和压感式。

3）数字化仪

在大量 GIS 应用中，一个至关重要的任务就是要把若干图形输入计算机中去。计算机不能直接识别这些图形，必须数字化即将其坐标输入到计算机。专门实现这一功能的计算机外部设备叫数字化仪。即所谓数字化仪是指专门用来读取图形信息的计算机输入装置。

数字化仪设备比较简单，一般由两部分组成，第一部分是感应板部分（又叫画图板 Drawing Board，但叫感应板比较确切），第二部分是点设备（Pointing devices）又叫传送器或者游标。对于立式的数字化仪还有一个底座，是为了架感应板用的。

数字化仪是计算机图形系统的输入设备。当画笔接触到其图形板上的某一位置时，会将画笔位置的坐标转换为二进制的数字量，输入计算机，随着画笔的运动，可以把一个图形输

入到计算机中。

数字化仪的种类较多，按测量坐标的原理，大体上可分为机电式、超声波式、磁致伸缩式和电磁感应式四种，其中电磁感应式是目前最常见的数字化仪。

4）鼠标器、跟踪球和操纵杆

鼠标器、跟踪球（也称轨迹球）和操纵杆，是与屏幕相配合拾取光标的相位坐标的输入设备。

1．4．5 虚拟现实设备

一个虚拟现实系统，可以分解为三个独立的、但又相互联系的感觉引导子系统，即视觉子系统、听觉子系统和触觉动觉子系统。这三个子系统由虚拟环境产生器进行控制、协调，如图 A1-3 所示。

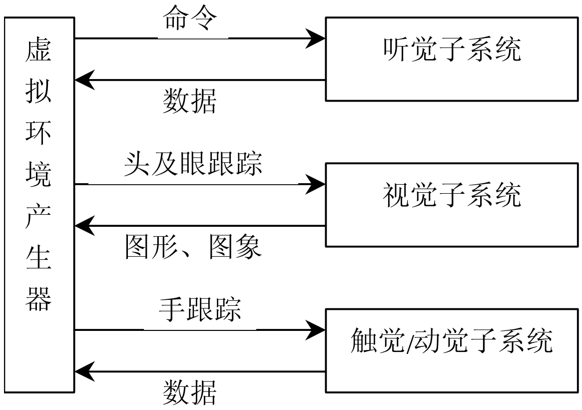


图 A1-3：VR 系统一般组成

1）虚拟环境产生器

虚拟环境产生器实质上是一个包括虚拟世界数据库的高性能计算机系统。该数据库包含了虚拟环境中对象的描述以及对象的运动、行为及碰撞作用等性质的描述。虚拟环境产生器的另一作用是生产图像。这些图像的生成必须在最短的时间延迟内考虑参与者头部的位置和方向。虚拟环境产生器内的任何通信延迟都必将表现为视觉的滞后。如果这种滞后可以感知，在某种条件下就会使参与者产生晕眩的感觉。

2）触觉/动觉子系统

为了增强虚拟环境中身临其境的感觉，必须给参与者提供一些诸如触觉等方面的生理反馈，触觉反馈是指 VR 系统必须提供所能接触到的物体的触觉刺激，如物体表面纹理或甚至包括触摸的感觉等。参与者感觉到物体的表面纹理等时，同时也感觉到运动阻力。当然，毫无疑问 VR 系统中的触觉/动觉反馈是很难实现的。一旦实现，将极大地增强虚拟存在的感受。目前触觉/动觉系统中一个重要的部分是手跟踪和手势跟踪。它的一个已经实用化的设备是数据手套（Data Glove），如图 A1-4 所示。

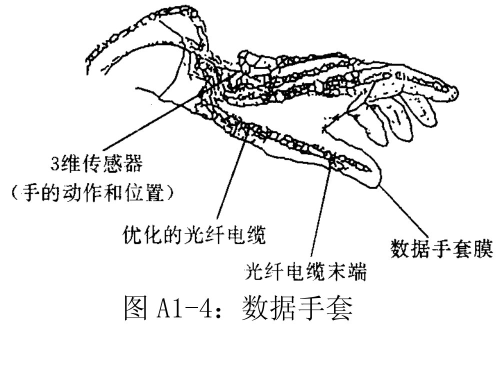


图 A1-4：数据手套

数据手套的机理主要依靠纤细的光导纤维和光线的直线传播特性。它选用非常适合于屈伸的材料制成。对每一个指头都有一根光纤从手腕出发，经指尖绕回再到手腕处的光纤；一端装有光信号源（LED），另一端装有测量光通量的光传感器件。在指关节处光纤表面切有微小的豁口，当手指弯曲时豁口裂开有光通量漏掉。当人带上手套后手指伸直时，由于光线的直线传播几乎能获得 100％的输出光量，一旦手指弯曲则光量随弯曲程度而衰减。这种光量的变化，在控制器里由模/数变换器（A/D）转换成数字量，向主计算机传送，并进行计算、解释。

目前，数据手套暂时只能输入手势语言信息，当人情不自禁地去“触摸”或抓放一个物体时，数据手套便可以把这些手势信息转入（反馈）到虚拟环境产生器中。当然，为了反映手在“抓摸”时的用力情况，还应有压力反馈，这个问题目前正在解决。

3）视觉子系统

视觉是人类用以接收信息的主要器官。目前，VR 技术中最重要的一项技术是大视场双眼体视显示技术。

人类的视觉，是一个具有双眼坐标定位功能自然序列：人的两只眼睛同时看到周围世界的同一个窗口，但由于两眼位置上的差别，在视网膜上各自生成略有差别的图像，这两个图像通过大脑，被综合成一个含有景物深度的立体图像。VR 体视显示技术用以下两种方案解决这一问题：一种是用两套主机分别计算并驱动对应左右眼的两个显示器；另一种是用一套主机分时地为左右两眼产生相应的图像。

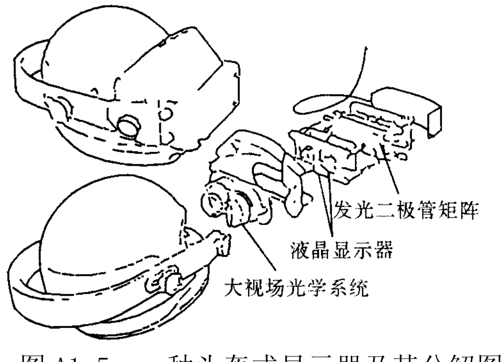


图 A1-5：一种头盔式显示器及其分解图

目前，VR 显示装置的主流是头盔式显示器。图 A1-5 为头盔式显示器原理的分解图。当然，最重要的还是要能在显示屏上产生清晰、逼真的图像。

4）听觉子系统

通常听觉系统也安装在头盔显示器上。听觉子系统主要由声音合成、3D 声音定域和语音识别组成，以给虚拟环境中的用户一个真实的声音环境。

（1）声音合成

尽管听觉系统以比视觉系统低得多的频带宽度工作，但人的听觉系统很善于在众多的声音中挑取特定的声音，作为对视觉摄取信息的补充。因此，在 VR 系统中加入声音合成装置是十分必要的。当视觉系统处理某一事件时，听觉系统同时在后台工作。

（2）3D 声音定域

为造成逼真的声音环境，就要使参与者能通过两耳因位置不同，所接受的声波的时差等，分辨出声源与自己的相对位置；即使参与者在头部运动时，也能感觉这种声音保持在原处不变。为了达到这种效果，声音定域系统必须考虑参与者两个“耳廓”的高频滤波特性。参与者头部的方向对于正确地空间化声音信号是很重要的。因此，虚拟环境产生器要为声音定域装置提供头部的位置和方向信号。

（3）语音识别

语音识别在输入数据大量时，是非常有效的。

1．4．6 调制解调器

目前使用的计算机一般都是数字计算机，即在计算机中处理的是数字信号，而普通电话线上传输的是音频信号。用普通电话线传输数字信号的效率是很低的。为了能用普通电话线进行计算机通信，应当把要发送的数字信号先调制（Modulate）成音频信号，送到目的地后再解调（Demodulate）成数字信号。完成这一功能的设备称为调制解调器 MODEM。由于一台计算机既要接收信号，又要发送信号，所以调制解调器既有调制功能，又有解调功能。调制解调器是拨号接入方式下的关键设备。

一些新的输入输出设备：

随着计算机技术的发展，出现了一些新的、更加方便的输入输出设备，下面是其中的简单介绍：

源数据自动化设备：包括条形码阅读机，磁性墨水字符阅读机（用于支票上的数字）。

语音录入：也称为“语音识别”，允许用户通过讲话向系统发出指令。

数码相机：可以直接得到数字图像，并由软件进行进一步的处理。

1．5 计算机系统性能

全面衡量一台计算机的性能要考虑多种指标，并且对不同的用途所侧重的方面不同。下面从普遍应用的角度，介绍主要的几种性能指标。

1）CPU 字长

CPU 字长是指 CPU 一次所能处理的位数。CPU 字长越长，表明 CPU 所能处理的数据的精度越高，并且影响处理的速度。因为短字长的 CPU 对较大的数据要通过两次甚至多次运算实现。目前微型计算机的字长从 8 位、16 位、32 位，到 64 位等。当然 CPU 字长越长，价格就越高。为了适应不同的需要，并协调精度与成本，人们还设计了可变字长计算，如半字长、全字长、双字长等。

2）主频率

CPU 工作的节拍是由主时钟控制的。主时钟不断地产生固定频率的时钟脉冲，时钟脉冲的频率就是 CPU 的主频率。主频率越高，CPU 的工作节拍越快。这是影响机器运算速度的重要参数。

3）主存容量

主存用以直接与 CPU 交换信息。主存容量大，处理问题的能力就强。同时由于它与外存之间的信息交换次数少，解题时间效率高。计算机的最大主存容量由 CPU 的地址总线的根数决定。地址总线为 16 条时，CPU 的最大寻址范围为 64K；地址总线为 20 条时，CPU 的最大寻址范围为 1M。目前，微机地址总线为 32 条，最大寻址空间为 4G。

4）软、硬件配置及性能价格比

软、硬件配置包括外部设备的配备情况，指令系统以及操作系统功能的强弱、界面是否友好，有无其它支持软件和应用软件等。性能价格比是人们对经济效益的选择，这个值越大越好。

5）RASIS 特性

可靠性（Reliability）、可用性（Availability）、可维护性（Servicebility）、完整性（Integrality）和安全性（Security）统称 RASIS。它们是衡量一个现代化的计算机系

统性能的五大功能特性。

6）兼容性

所谓兼容性（Compatibility），是指系统间所含的某些“东西”具有并存性，即意味着两个系统之间存在着一定程度的通用性，它使机器能承前启后、便于推广。

2．数据库系统基础

目前，数据库管理系统（DBMS,Data Base Management System）正日益进入最终用户的日常应用，人们每天都在日常生活中用到数据库，如使用信用卡购物、订票、书目查询等等，在使用过程中，用户不需要了解数据的具体存取和管理方式，正是数据管理系统提供了这些功能。

2．1 数据库的基本概念

定义数据库管理系统之前，必须首先定义这种系统的基本成分——数据库，一个数据库有四个主要成分：数据、联系（Relationship）、约束（Constraint）和模式（Schema）（图 A1-6）。数据是所存储的逻辑实体在计算机中的二进制表示；联系表示数据项之间的某种对应；约束是定义正确数据状态的断言；一种模式描述数据库中数据的组织和联系。

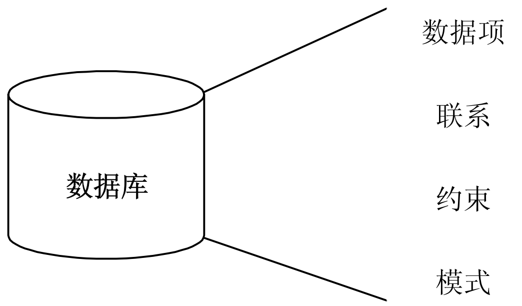


图 A1-6：数据库组成

模式为数据库管理系统各个组成部分的使用和应用的安全定义数据库的各种视图。模式将数据存储的物理外表与逻辑表示分开（见图 A1-7）。内部模式定义数据在物理数据存储区中如何组织以及放在何处。概念模式模型按照适当的数据库数据模型（如关系模型或对象模型）定义所存储数据的结构。外部模式为特定用户们定义数据库的一个或多个视图。

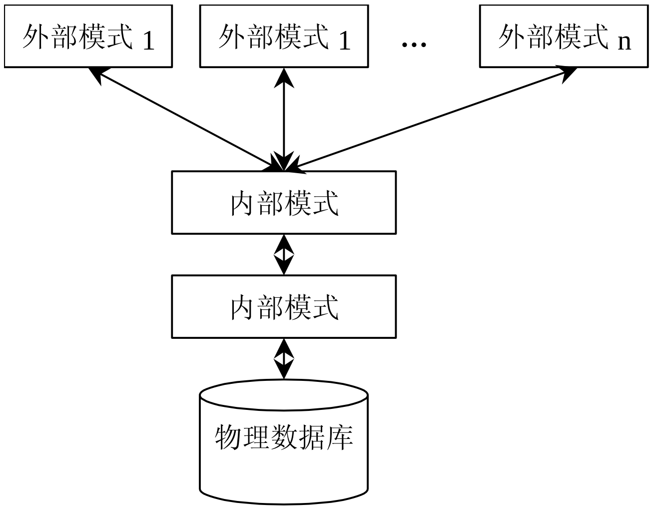


图 A1-7：数据库模式的概念

数据库管理系统为访问数据库提供服务，同时维护存储数据所要求的正确性和一致性。

在数据库管理系统中，运行的工作单元是事务（Transaction）单元，在此之上定义了一致性和正确性。事务应该支持 ACID 属性。ACID 属性包括：事务运行的原子性（Atomic）、一致性（Consistent）、独立性（Isolation）和事务执行的持久性（Durability）。

1）原子性确保事务被当成一个不可分割（All－or－Nothing，要么全做，要么根本不做）的操作单元处理。

2）事务操作的一致性确保数据库从原来的一致性状态正确转移到一个新的一致性状态，此处的一致性由数据库中数据项上的谓词定义。

3）事务的独立属性定义了允许它们可见什么。一个被孤立的事务只“看到”数据库的一个视图，就好像事务是单独在数据库上执行一样。

4）持久性属性确保一旦提交了一个事务，其结果就持久存在且不能从数据库中消除。

2.2 数据库的数据模型

按照描述数据与数据间关系的方法不同，数据库常用的数据模型有层次模型、网状模型和关系模型。

2．2．1 层次模型

层次模型是一种基本层次联系的集合，它实际上是一种有根定向的有序树，如图 A1-8 所示。层次模型的基本结构是树结构——根、枝、叶结构，数据存放的基本单位是片断（即层），片断是内在有逻辑联系的一组数据，总的来说，层次模型按照树形结构以片断为单位存放数据。层次模型比较容易实现，但是查找比较麻烦，数据的冗余度也比较大。

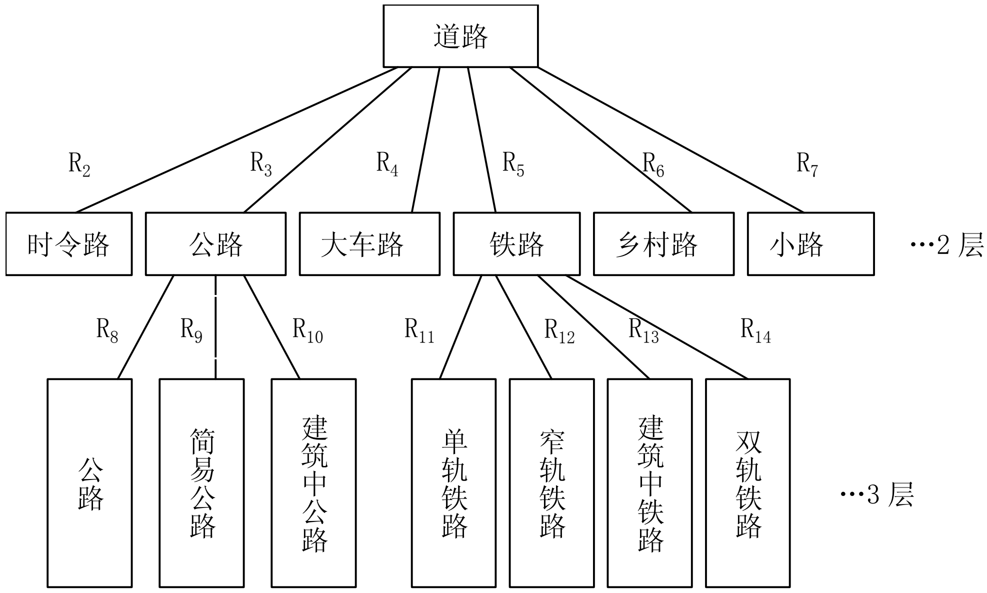


图 A1-8：数据库层次模型

2．2．2 网状模型

所谓网状模型是指一个连通的基本层次联系的集合，如图 A1-9 所示。复杂的网总可以分解成若干个基本结构，即分解成系。系有系主（只有一个）和系属（可以有多个），系主和系属之间有关系，而且关系是双向的。

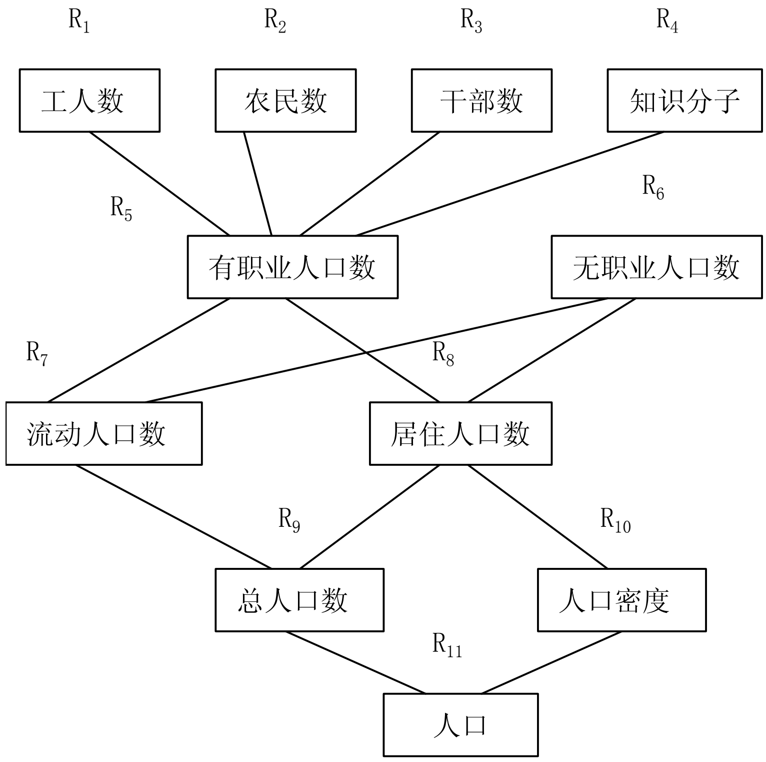


图 A1-9：网状模型示例

网状模型存放的基本单位是记录，也就是按记录存放，查询时，从系（系主和系属）查起。网状模型查找时间较省，数据和冗余度比层次模型小，但比关系模型要大。

2．2．3 关系模型

关系模型是目前最为流行的数据模型，它是由许多二维关系表组成的集合。例如表 1 就是一张关系表，R是关系名，A_i是属性名，关系和属性（R,A_1,A_2…A_n）组成了数据表的模式(Schema)；V_{ij}叫做分量，表中的一列是一个属性，相当于一个数据项（或数据元素），一行叫做一个元组（Tuple），相当于一条记录。

关系模型中所有数据都按表格存放，有关系的数据放在一张表上，表与表之间有连接。

表 A1-1：关系表 R

A_1	A_2	… … …	A_j	… … …	A_n
V_{11}	V_{12}	… … …	V_{lj}	… … …	V_{ln}
：：： V_{il} ：：：	：：： V_{i2} ：：：	：：： ：：：	：：： V_{ij} ：：：	：：： ：：：	：：： V_{in} ：：：
V_{m1}	V_{m2}	… … …	V_{mj}	… … …	V_{mn}

关系模型的特点是每一个分量必须是不可分割的数（数据元素），也就是不允许表中还有表。另外表和表之间的联系要用关系表示，而不是用其它表示。

关系模型查找很方便，数据冗余度小，但关系联结时效率较低。

关系表的操作可以分为以下四种：

1）通用的集合操作，如并、交、差运算等；

2）去除关系表的某些部分的操作，包括选择(Selection)和投影(Projection)，前者去除某些元组，后者则用于除去某些属性；

3）两个关系表的合并，包括“笛卡尔积”以及各种方式的连接运算；

4）更名操作，即对关系表属性名称的修改，它不改变元组，但是改变了关系表的模式。

下面以常见的学生管理例子说明这些操作：

学生关系表 R1

姓名	年龄	性别
Carola	20	女
Maxwell	21	男
Churchill	18	男

学生关系表 R2

姓名	年龄	性别
Johnson	17	男
Maxwell	21	男
Barbara	24	女
Williams	20	女

学生成绩表 R3

姓名	语文	数学
Johnson	A	A
Maxwell	B	A
Barbara	A	A
Williams	A	C
Carola	C	C
Churchill	C	B

对 R1 和 R2 进行求交集运算的结果（R4）

姓名	年龄	性别
Maxwell	21	男

对 R1 和 R2 进行求并集运算的结果（R5）

姓名	年龄	性别
Johnson	17	男
Maxwell	21	男
Barbara	24	女
Williams	20	女
Carola	20	女
Churchill	18	男

对 R5 和 R3 根据“姓名”进行连接操作结果（R6）

姓名	年龄	性别	语文	数学
Johnson	17	男	A	A
Maxwell	21	男	B	A
Barbara	24	女	A	A
Williams	20	女	A	C
Carola	20	女	C	C
Churchill	18	男	C	B

对 R6 选择“性别=“男””，并且对姓名、语文、数学成绩进行投影的结果（R7）

姓名	语文	数学
Johnson	A	A
Maxwell	B	A
Churchill	C	B

关系数据库的查询和修改操作是通过 SQL(结构化查询语言，Structured Query Language)，SQL 的基础是关系代数，同时也包括了其它的一些操作，如求和、汇总、数据更新等等，下面是一个 SQL 查询的例子：

对于关系：

省份（省会，人口，产值，面积）

SQL 语句：

*Select **

From 省份

Where 人口>60,000,000 and 面积<160,000

用于查询得到人口大于 6 千万并且面积小于 16 万平方公里的省份。

2．2．4 面向对象数据库模型

网络和层次以及关系模型都适合那些结构简单以及访问有规律的数据。这些模型的最佳应用领域有个人记录管理，清单控制，终端用户销售，商业记录等，所有这些应用领域都只有相当简单的数据结构、联系以及数据使用模式。但是，当试图把这三种模型应用于更高级的领域时，数据不能用类似于记录这样的简单结构来表示了，访问和操作方法也不再简单。这些应用领域需要更复杂的抽象数据类型，如图形，声音，图标，包，清单，队列，以及地图，这些数据类型都各自定义了独特的操作方法——例如，一个地图对象可以定义为经度、纬度、地点的时间维；以等高线来定义地形；用图标表示主要的嵌入对象，而他们本身也可能是对象。除了这些定义之外，在地图的各区域可能还含有隐藏的数据。我们可以表示人口密度、动物密度、植物、水源、建筑物及其类别（例如，单个住宅楼，高楼，工业建筑，居民楼）、污染情况以及其他信息，所有这些都是从应用领域典型使用中派生出来的抽象数据类型（图 A1-10）。

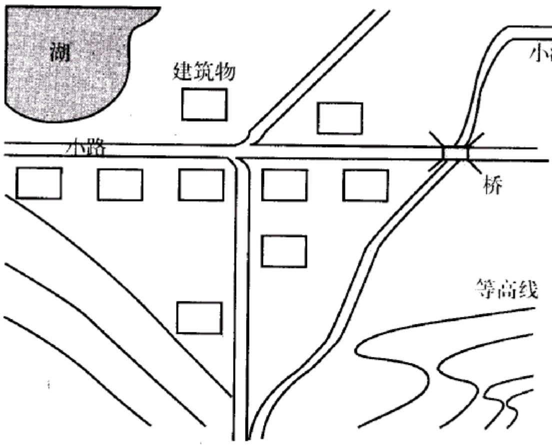


图 A1-10：地图抽象数据类型

面向对象数据库的引入就是为了满足一再出现的复杂信息的共享。在复杂数据进入数据库以后，数据库提供了存贮信息的统一视图，与具体存贮结构无关。把物理数据结构与逻辑数据结构分开，同时控制数据的共享及保持数据的正确性、完整性和一致性，大大方便了应用程序的开发和维护，减少生命周期内的各种费用。通过一组优化的程序来管理数据，使得整体效果更优，性能更稳定。

2.3 数据库管理系统

数据库管理系统（DBMS）是为数据库访问提供服务的软件，同时维护所有数据必需的特性。数据库管理系统为支持应用程序访问和操作数据库数据提供下列服务：

1）事务处理

事务将使数据库从一个一致状态转移到另一个一致状态。数据库操作被分成两大类：数据访问操作和事务操作。有三种特定的事务操作：启动（Start）指示将开始一个新事务，提交（Commit）指示事务已正常终止且其作用结果将持久存在，以及放弃（Abort）指示事务被异常终止，其所有结果将被放弃。事务通常需要具有前面提到过的 ACID 属性。

2）并发控制

并发控制是一种数据库管理活动，它协调数据库操作进程的并发操作和对共享数据的访问，并且解决它们之间可能发生的潜在冲突。并发控制机制的目标是允许并发维护共享数据的一致性，数据库系统中的并发单元是事务。

3）恢复

数据库中恢复的目标是确保异常终止或出错的事务不会对数据库或其他事务产生不利影响。异常终止的事务有两种影响；对数据的影响和对其他事务的影响。恢复可使得数据库在事务异常终止后返回某个一致状态。

4）安全

安全是保护数据免受非授权的泄露、更改或破坏。每个用户和应用程序都有特定的数据访问特权。这些特权可以由外部模式定义，即根据各个用户被允许访问和/或修改的数据，给予它们不同的数据视图。安全系统提供一些方法，来决定每个用户或应用程序可访问什么视图。通过授权和身份鉴别过程，安全还具有限制初始访问数据库的功能。这些过程中最常用的是注册名和口令保护服务。

5）语言接口

DBMS 提供对用于定义和操作数据的语言的支持。概念模式是用数据定义语言（Data

Definition Language，DDL）说明的。这种数据库语言部件是用来描述数据、数据间联系和对数据和联系的约束的表示法。DDL 首先用在数据库设计时，以后修改模式时还会用到 DDL。

数据操纵语言（Data Manipulation Language，DML）用于表达数据库上的操作。DML 有时也称为查询语言。DBMS 提供 DML，以便用户和应用程序编写者访问数据库中数据，而不必知道数据库如何存储数据或把数据存在何处。

6)容错性

不管发生什么故障仍能继续提供可靠 DBMS 服务的能力称为容错性。一个出错的数据库部件将使与其交互的其他部件产生故障。典型的数据库故障包括违反约束和事务超时错误。如上所述，恢复与容错性密切相关，因为恢复是一种机制，它能容许发生使事务异常终上的差错。

7)数据目录

数据目录（有时称为数据字典）是一个系统数据库，它含有主数据库中数据的描述（有时被称为元数据，Metadata）。它包含有关数据、联系、约束的信息，以及将这些特征组织到一个统一数据库中的所有模式的信息。通过查询目录可获得有关主要数据库结构的信息，因而目录被看成一个数据库。

8）存储管理

DBMS 提供数据持久存储的管理机制。内部模式定义数据应该如何用存储管理机制存储。为了访问物理存储，存储管理系统与操作系统间有接口。

3．数据结构和算法

3.1 数据结构的概念

3．1．1 数据结构——一个例子

从提出一个实际问题到计算机解出答案需要经过下列步骤：首先从实际问题抽象出一个数学模型，然后设计一个解此数学模型的算法，最后编出程序、进行测试、调整直至得到最终解答。寻求数学模型的实质是分析问题，从中提取操作的对象以及这些操作对象之间含有的关系，然后用数学语言加以描述。例如，在分析了一个物理现象或化学现象变化的规律之后可以得到一组代数方程或微分方程。然而，更多的问题无法用数学方程加以描述。下面用一个例子来说明数据结构。

多叉路口交通灯的管理问题。通常，在十字交叉的路口只要设红绿两色的交通灯便可保持正常的交通秩序，而在多叉路口需设几种颜色的交通灯才能既使车辆相互之间不碰撞而又达到最大的流通呢？假设有如图 A1-11（a）所示的五叉路，其中 C 和 E 为单行道，在路口有 13 条可行的通路，其中有的可以同时通行，如 A→B、和 E→C，而有的不可同时通行，如 E→B 和 A→D，那末，在路口应如何设置交通灯进行管理？这个问题可以转换成一个图的染色问题。假设在图上以一个圆圈表示一条通路，在不能同时通行的两个圆圈之间画一连线，对图中的圆圈上色，要求同一连线上的两个圆圈不同色且颜色种类最少。图 A1-11（b）是与图 A1-11（a）相应的图，图中 13 个圆圈表示 13 条通路，圆中的号码分别表示四种颜色的交通灯。

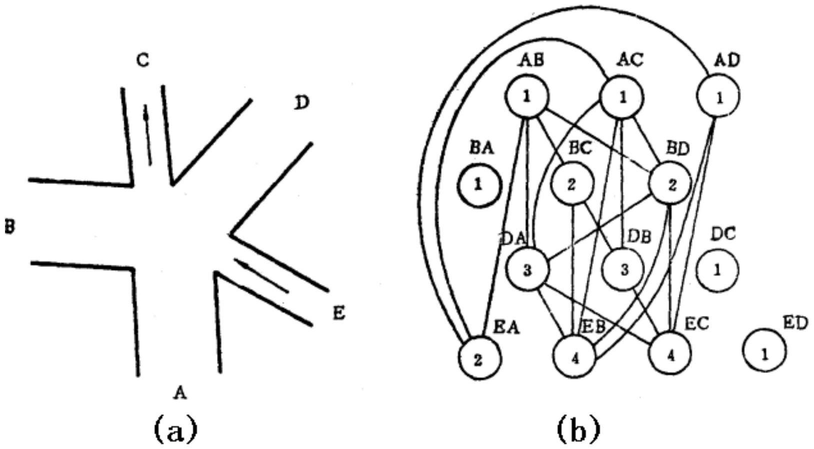


图 11:五叉路口交通管理意图

（a）三叉路口（b）表示通路的图

由以上例子可见，描述这样一类问题的数学模型不再是数值方程，而是诸如表、树和图等的非数值性的数据结构。因此，简单说来，数据结构就是一门研究非数值性程序设计中计算机操作的对象以及它们之间的关系和运算等的学科。

3．1．2 基本术语

下面是数据结构中常用到的名词和术语的含义：

1）数据（Data）

数据是描述客观事物的数、字符以及所有能输入到计算机中并被计算机程序处理的符号的集合。它是计算机程序加工的“原料”。例如，一个利用数值分析的方法解代数方程的程序处理的对象只是整数和实数，而一个编译程序或文字处理程序的对象是字符串。因此，对计算机而言，数据的含义极为广泛，如图形、声音等都属于数据的范畴。

2）数据元素（Data Element）

是数据的基本单位，即数据这个集合中的一个个体（客体）。有时一个数据元素可由若干个数据项（data item）组成，数据项是数据的最小单位。

3）数据对象（Data Object）

是具有相同特性的数据元素的集合，是数据的一个子集。例如，整数的数据对象是集合 $N=\{\cdots,-2,-1,0,1,2,\cdots\}$，字母字符的数据对象是集合 $C=\{A,B,\cdots,Z\}$。

4）数据结构（Data Structure）

简单说来，数据结构是带有结构的数据元素的集合。从上述面的例子可以看到，被计算机加工的数据元素都不是孤立的，在它们之间存在着某种联系，这种相互的关系，通常称做结构。我们可以从集合论的观点加以形式化描述。

数据结构是一个二元组

Data-Structure＝（*D,R*）

其中：D 是数据元素的集合，R 是 D 上关系的集合。例如，复数可被定义为一种数据结构:

Complex＝（*D, R*）

其中 D＝｛$x|x$ 是实数｝

$R=\{R1\}$

$R=\{\langle x,y\rangle / x,y\in D$，$x$ 称为实部，y 称为虚部$\}$

由此，由任何一对实数均可得到一个复数。

在上面对数据结构的定义中，R 集合中的关系指的是数据元素之间的逻辑关系，因此，又称数据的逻辑结构。与此对应，则有数据的物理结构，又称存贮结构，是数据结构在计算机中的映象（或表示）。

5）数据类型（Data Type）

是程序设计语言中所允许的变量的种类，换句话说，是变量可能取的值和能作的运算的集合。在每一种程序设计语言中都有一组它所允许的基本数据类型。各种语言所能提供变量类型的多少决定了该语言功能的强弱。我们可以把数据类型看作是程序设计语言中已经实现的数据结构，如复数、数组等。因此，数据类型实际上是数据结构（包括逻辑结构和存贮结构）及其运算的总称。

3.2 两种重要的数据——树和图

3．2．1 树和二叉树

树（Tree）是一种重要的非线性结构，在计算机软件设计中被广泛使用，如哈夫曼树可以应用于数据压缩编码，而B^+树则在文件系统管理中被使用。图A1-12 表示了树的示例。

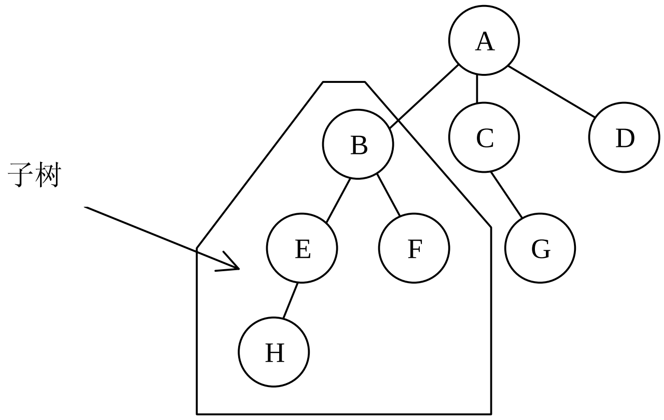


图 A1-12：树的示例

在图 A1-12 中，结点 A 为树的根，根的每个分支称为子树（Subtree），子树也是一棵树；结点子树的根为结点的孩子（Child），如 B、C、D 为结点 A 的孩子，而 A 为 B、C、D 的双亲（Parent）；同一个双亲的孩子之间为兄弟（Sibling）关系；没有孩子的结点为树的叶子（Leaf），H、F、G、D 为树的叶子。

树的基本操作包括：

1）初始化一棵树；

2）得到树的根；

3）得到一个结点的双亲；

4）得到一个结点的兄弟；

5）得到一个结点的孩子；

6）插入子树；

7）删除子树；

8）遍历（Travers）树；

9）清空树。

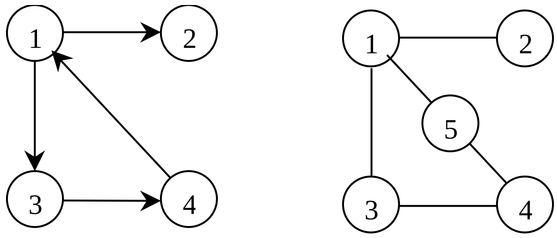
在不同的软件系统中，由于具体实现的不同树的操作也不同。

二叉树（Binary Tree）是另一种树型结构，其特点是每个结点最多有两棵子树。树和二叉树之间可以相互转换。

各种特定的树型结构被广泛应用于查找算法的实现中，它们可以加快查找的速度；在地理信息系统中，可以用于空间索引的建立，以提高空间要素的检索效率。

3．2．1 图

图（Graph）是较树更为复杂的数据结构，在图中，结点之间的关系是任意的，图中任意两个数据元素都可能相关。图 A1-13 给出了图的示例。



（a）有向图　　　　　　（b）无向图

图 A1-13：图的示例

图的形式化定义为：

$Graph = (V,R)$

其中：

$V=\{x|x \in dataobject\}$

$R=\{VR\}$

$VR=\{<x,y>|P(x,y) \wedge (x,y \in V)\}$

图中数据元素称为顶点（Vertex），V 是顶点的有穷非空集合；VR 是两个顶点之间的关系的集合，其定义中 *P(x,y)*表示 x 到 y 的一条单向通路；若$<x,y> \in VR$，则$<x,y>$表示从 x 到 y 的一条弧，此时图称为有向图（Digraph）；若$<x,y> \in VR$ 必有$<x,y> \in VR$，则此时图称为无向图。

在图中，如果$<x,y> \in VR$，则x，y互为邻接点。路径（Path）是一个顶点序列（$V_1, V_2, \ldots V_n$），其中V_i和V_{i+1}为邻接点。

图可以有多种存储结构，其中最普通的是采用邻接矩阵，如果两个结点$<V_i, V_j> \in VR$，则矩阵对应元素A[i,j]=1，反之，A[i,j]=0。

无向图的邻接矩阵是对称的，而有向图的邻接矩阵则不一定对称，图 A1-13 中两个图的邻接矩阵如下（其中没有考虑相同点邻接性）：

$$a=\begin{bmatrix}1&0&1&0\\0&0&0&0\\0&0&0&1\\1&0&0&0\end{bmatrix}$$

$$b=\begin{bmatrix}0&1&1&0&1\\1&0&0&0&0\\1&0&0&1&0\\0&0&1&0&1\\1&0&0&1&0\end{bmatrix}$$

在地理信息系统中，可以应用图数据结构进行网络分析，判断两个空间要素之间的连通性及其最短路径。

3.3 算法

3．2．1 算法的概念和特性

数据结构和算法构成了计算机程序，所谓算法（Algorithm），是对特定问题求解步骤的一种描述，它是指令的有限序列，其中每个指令表示一个或者多个操作。一个算法还具有以下五个重要特征：

1）有穷性

一个算法必须总在执行有穷步之后结束，而且每一步都可以在有穷时间内完成；

2）确定性

算法中的每一步必须有明确的含义，不会引起二义性；并且，算法只有唯一的执行路径，即相同的输入只会得到相同的输出。

3）可行性

一个算法必须是可行的，即算法中的操作都是可以通过已有的基本运算完成。

4）输入

一个算法可以有零个或者多个的输入。

5）输出

一个算法有一个或者多个的输出，这些输出同输入有特定的关系。

3．2．2 算法设计的要求

设计一个“好”的算法应该考虑达到以下目标：

1）正确性

算法应该满足具体问题的要求，对于输入数据能够得到符合规格说明的结果。

2）可读性

算法的可读性有助于人对算法的理解，便于修改和调试。

3）健壮性

当输入数据非法时，算法也能够正确处理，而不会出错。

4）效率与低存储量需求

效率指的是算法执行时间，一个算法执行时间越少，效率越高；存储量需求是指算法执行过程中所需要的最大存储空间。一个好算法要有高的执行效率和低的存储量需求，但是在实际实现中，二者并不可以兼得，通常可以用大的空间开销来提高算法执行效率，或者以降低执行速度作为减少空间开销的代价。

决定一个算法的效率和存储需求的因素包括书写程序的语言、生成的机器代码的质量、机器执行指令的速度等等，但最重要的是算法解决问题的规模。一个算法是由控制结构（顺序、分支和循环三种）和原操作构成，通常是在算法中选取一个基本原操作，并以该操作重复执行的次数作为算法的时间度量。

一般情况下，算法中基本操作重复执行的次数是问题规模 n 的某个函数 f(n)，算法的时间量度记做：

T(n)=O(f(n))

它表示随问题规模n的增大，算法执行时间的增长率和f(n)的增长率相同，称为算法的渐近时间复杂度，简称时间复杂度(Time Complexity)。常见的算法时间复杂度有O(1)、O(n)、平方阶O(n^2)、对数阶O(log$n$)、指数阶O(2^n)等等，在算法设计时，可以根据时间复杂度对算法优劣进行评判，通常指数阶的时间复杂度是不可以考虑的。

算法执行所需要的存储空间通过空间复杂度来度量，记做：

S(n)=O(f(n))

其意义类同于时间复杂度。

软件，是“固化”的思想。

杨芙清

附录二 GIS 工具软件介绍

导读：本章介绍了比较常用的一些地理信息系统软件，具体包括三家美国 GIS 开发商 ESRI，Intergraph 和 MapInfo 的软件产品，以及三个国产软件：MapGIS，GeoStar 和 Citystar。

介绍的内容来自于各个公司的宣传材料以及 WWW 上各个公司的主页，但是除去了诸如“第一”，“最好”等的修饰词汇。另外，由于软件在不断升级，在本章中也不具体说明软件版本号，并且也不评判软件的优缺点，只是简要列出了各个软件的模块名称和功能目标。

这些软件，提供了相似的功能集合，不同之处在于其具体的实现方式（如用户界面，操作流程），和操作效率（如速度、数据量）。这种相似性，也正说明了 GIS 技术的成熟。

1．ESRI 产品系列

ESRI 公司（Environmental Systems Research Institute Inc.）于 1969 年成立于美国加利福尼亚州的 Redlands 市，公司主要从事 GIS 工具软件的开发和 GIS 数据生产。

ESRI 的产品中，最主要的是运行于 UNIX/Windows NT 平台上的 ArcInfo，它由两部分组成：Workstation ArcInfo 和 Desktop ArcInfo。

1）Workstation ArcInfo 基于拓扑数据模型，实现了图库（Map Library）的管理，并且具有了栅格数据的分析功能，支持栅格矢量一体化查询和叠加显示。此外，ArcInfo 还提供了二次开发语言 AML 以及开放开发环境 ODE，以便于用户定制自己的 GIS 应用。

Workstation ArcInfo 提供了最基本的 GIS 功能，包括数据录入和编辑、投影变换、制图输出、查询分析及其分析功能（缓冲区分析、叠加复合分析等等）。

除了上述基本功能以外，Workstation ArcInfo 还通过一些扩展模块实现特定的专门功能：

TIN：基于不规则三角网的地表模型生成、显示和分析模块，可以根据等高线、高程点、地形线生成 DEM，并进行通视、剖面、填挖方计算等。

GRID：栅格分析处理模块，可以对栅格数据进行输入、编辑、显示、分析、输出，其分析模型包括基于栅格的市场分析、走廊分析、扩散模型等。

NETWORK：网络分析模块，提供了最短路径选择、资源分配、辖区规划、网络流量等功能，可以应用于交通、市政、电力等领域的管理和规划。

ARCSCAN：扫描矢量化模块。

ARCSTORM：基于客户机/服务器机制建立的数据库管理模块，可以管理大量的图库数据。

COGO：侧重于处理一些空间要素的几何关系，用于数字测量和工程制图。

ArcPress：图形输出模块，可以将制图数据转换成为 PostScript 格式，并可分色制版。

ArcSDE：SDE 指空间数据引擎（Spatial Database Engine），它是一个连续的空间数据模型，通过它可以将空间数据加入到关系数据库管理系统中去，并基于客户机/服务器机制提供了对数据进行操作的访问接口，支持多用户、事物处理和版本管理。用户可以以 ArcSDE 作为服务器，定制开发具体的应用系统。

ARC/INFO 的图库管理

为了能够管理分布在不同图幅的多个专题要素，在 ARC/INFO 的图库中，把地图数据纵向分为“图层（Layer）”，而水平方向分为“图块（Tile）”，如图 A2-1 所示。

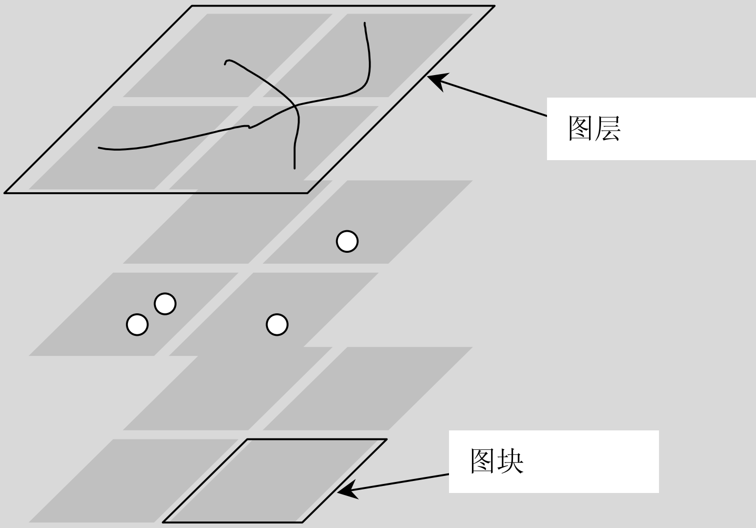


图 A2-1：ARC/INFO 的图库管理

在上图中，描述同一区域的不同专题图块构成一个“地图（Map）”，基于这种方式的管理，可以根据内容或区域范围任意调入相关的数据，并且便于实现数据共享，和并发访问控制。

2）Desktop ArcInfo 包括三个应用：Arc Map、Arc Catalog 和 Arc Toolbox。Arc Map 实现了地图数据的显示、查询和分析；Arc Catalog 用于基于元数据的定位、浏览和管理空间数据；Arc Toolbox 是由常用数据分析处理功能组成的工具箱。

3）ArcView GIS（图 A2-2）是 ESRI 的桌面 GIS 系统，它以工程为中心，实现了对地图数据、结构化的属性数据、统计图、地图图面配置、开发语言等多种文档的管理。除了提供脚本语言 Avenue 使用户可以定制系统以外，ArcView 还以“插件”的形式提供了一些扩展模块，包括：

Spatial Analyst：栅格数据的建模分析；

Network Analyst：网络分析；

ArcPress：制图输出；

3D Analyst：利用 DEM 实现三维透视图的生成；

Image Analyst：影像分析处理；

Tracking Analyst：通过直接接收、回放实时数据，实现对 GPS 的支持。

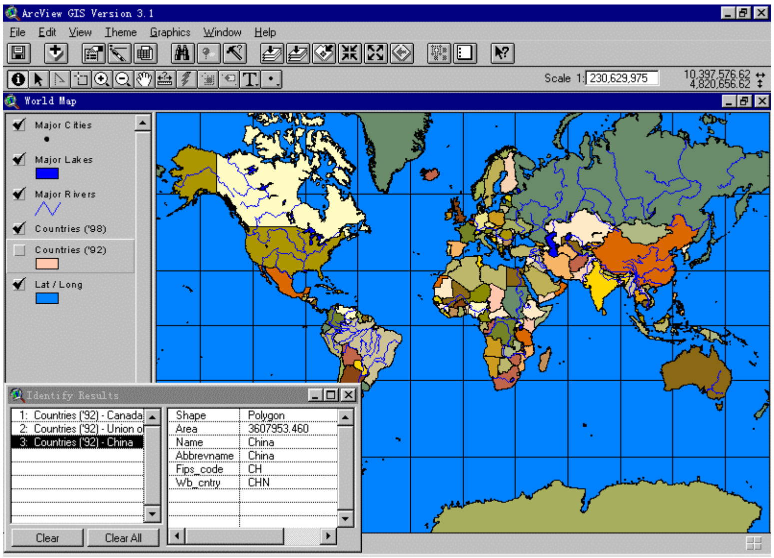


图 A2-2：ArcView GIS 用户界面

4）MapObjects 是一组供应用开发人员使用的 GIS 功能 OCX（OLE Custom Control）控件，用户可以采用其它的支持 OCX 的开发平台，如 Visual Basic，Delphi 等，集成 MapObjects，建立具体的应用系统。

5）ArcFM，支持公共设施规划、管理和服务的模块。

6）Internet Map Server(IMS)，实现了因特网上地理数据发布功能。

2．Intergraph 产品系列

Intergraph 公司成立于 1969 年，总部位于美国阿拉巴马州的汉斯维尔市，公司致力于计算机辅助设计、制造以及专业制图领域的硬件软件以及服务支持。

Intergraph 提供的 GIS 产品包括专业 GIS 系统（MGE），桌面 GIS 系统（GeoMedia），以及因特网 GIS 系统（GeoMedia Web Map）。

1）MGE 构成了 Intergraph 专业 GIS 软件产品族，它包括多个产品模块，提供了从扫描图像矢量化（I/GEOVEC），拓扑空间分析（MGE Analyst）到地图整饰输出（MGE Map Finisher）的基本 GIS 功能，此外还包括了其它一些扩展模块，实现了图像处理分析（I/RAS C，MGE Image Analyst），网络分析（MGE Network Analyst），格网分析（MGE Grid Analyst），地形模型分析（MGE Terrain Analyst），基于真三维的地下体分析（MGE Voxel Analyst）等一系列增强功能。

2）GeoMedia Professional（图 A2-3）设计成为与标准关系数据库一起工作，用于空间数据采集和管理的 GIS 产品，它将空间图形数据和属性数据都存放于标准关系数据库（Microsoft Access）中，在一定程度上提高了系统的稳定性和开放性，并且提高了数据采集、编辑、分析的效率。它支持多种数据源，包括其它 GIS 软件厂商的数据文件以及多种关系数据库；实现了矢量栅格的集成操作；提供了多种空间分析功能；此外，GeoMedia 包含其它一些模块，以应用于不同的具体领域。

GeoMedia Network：可以应用于交通网络以及逻辑网络的管理、分析、规划，具体包括最短路径查询、线路规划等功能。

GeoMedia SmartSketch：具有较强的图形编辑能力，是一个计算机辅助设计(CAD)软件。

GeoMedia Relation Moduler：用于建立设备间的网络关系，可以应用于自来水，煤气等市政管网的管理以及设备跟踪。

GeoMedia Object：GeoMedia 是基于控件的系统，它包含多个 OCX 控件，基于这些控件，用户可以开发具体的应用系统。

GeoMedia MFworks：基于栅格数据的分析模块，包含多种控件操作函数。

GeoMedia Oracle GDO Server：可以将地理数据写入到 Oracle 数据库并读出。

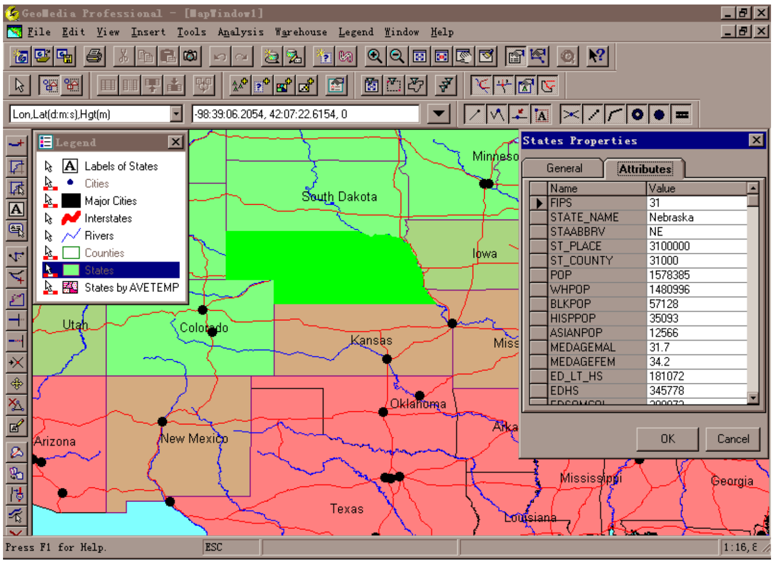


图 A2-3：GeoMedia 用户界面

3）GeoMedia WebMap 是 Intergraph 提供的基于因特网的空间信息发布工具。它提供了多源数据的直接访问和发布，并且支持多种浏览器。GeoMedia WebMap Enterprise 除了能够在因特网上发布数据之外，还提供了空间分析服务，如缓冲区分析、路径分析、地理编码等，用户可以在客户端通过浏览器提出请求，并输入具体参数，服务器进行计算并将结果返回给用户。

3．Mapinfo 产品系列

MapInfo 公司于 1986 年成立于美国特洛伊(Troy)市，成立以来，该公司一直致力于提供先进的数据可视化、信息地图化技术，其软件代表是桌面地图信息系统软件——MapInfo。

1）MapInfo Professional 是（图 A2-4）MapInfo 公司主要的软件产品，它支持多种本地或者远程数据库，较好地实现了数据可视化，生成各种专题地图。此外还能够进行一些空间查询和空间分析运算，如缓冲区等等，并通过动态图层支持 GPS 数据。

2）MapBasic 是为在 Mapinfo 平台上开发用户定制程序的编程语言，它使用与 BASIC 语言一致的函数和语句，便于用户掌握。通过 MapBasic 进行二次开发，能够扩展 MapInfo 功能，并与其它应用系统集成。

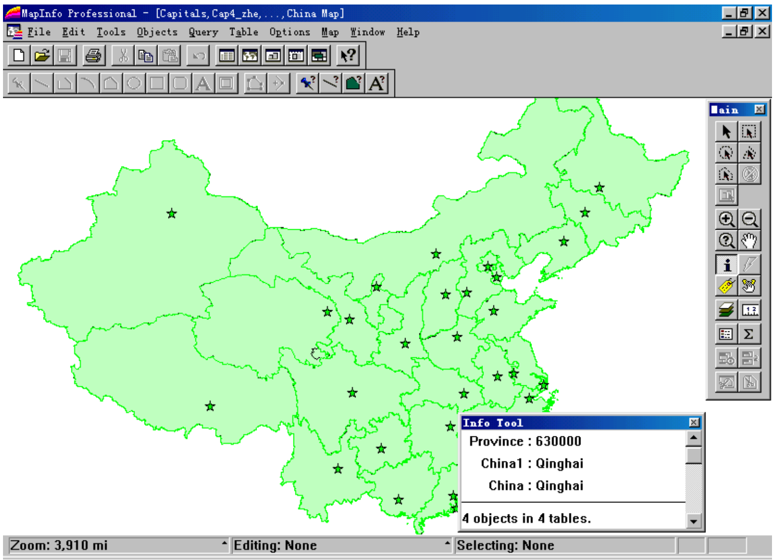


图 A2-4：MapInfo Professional 用户界面

3）MapInfo ProServer 是应用于网络环境下的地图应用服务器，它使得 MapInfo Professional 运行于服务器端，并能够响应用户的操作请求；而客户端可以使用任何标准的 Web 浏览器。由于在服务器上可以运行多个 MapInfo Professional 实例，以满足用户的服务请求，从而节省了投资。

4）MapInfo MapX 是 MapInfo 提供的 OCX 控件。

5）MapInfo MapXtrem 是基于 Internet/Extranet 的地图应用服务器，它可以用于帮助配置企业的 Internet。

6）SpatialWare 是在对象——关系数据库环境下基于 SQL 进行空间查询和分析的空间信息管理系统，在 SpatialWare 中，支持简单的空间对象，从而支持空间查询，并能产生新的几何对象。在实际应用中，一般使用 SpatialWare 作为数据服务器，而 MapInfo Professional 作为客户端，可以提高系统开发效率。

7）Vertical Mapper 提供了基于网格的数据分析工具。

4．国产 GIS 软件

4．1 MapGIS

MapGIS（图 A2-5）是中国地质大学开发的地理信息系统软件，其功能模块包括：

1）数据输入模块：提供了各种的空间数据输入手段，包括数字化仪输入，扫描矢量化输入以及 GPS 输入。

2）数据处理模块：可以对点、线、多边形等多种矢量数据进行处理，包括修改编辑、错误检查、投影变换等功能。

3）数据输出：可以将编排好的图形显示到屏幕或者输出到指定设备上，也可以生成 PostScript 或 EPS 文件。

4）数据转换：提供了 MapGIS 与其它系统之间数据转换的功能。

5）数据库管理：实现了对空间和属性数据库管理和维护。

6）空间分析：提供了包括 DTM 分析、空间叠加分析、网络分析等一系列空间分析功

能。

7）图像处理：图像配准镶嵌以及处理分析模块。

8）电子沙盘系统：实时生成地形三维曲面。

9）数字高程模型：可以根据离散高程点或者等高线插值生成网格化的 DEM，并进行相应的分析，如剖面分析、遮蔽角计算等等。

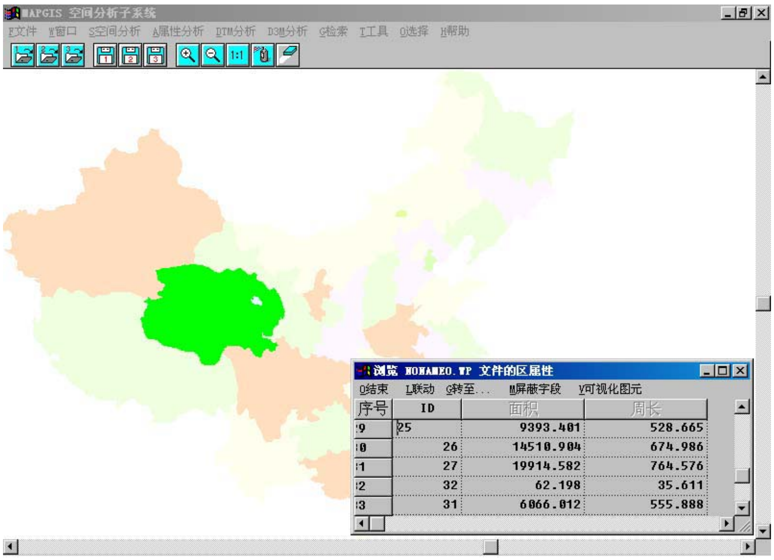


图 A2-5：MapGIS 用户界面

4．2 GeoStar

CeoStar（吉奥之星）是武汉测绘科技大学开发的、面向大型数据管理的地理信息系统软件，其功能模块包括：

1）GeoStar：是整个系统的基本模块（图 A2-6），提供的功能包括空间数据管理、数据采集、图形编辑、空间查询分析、专题制图和符号设计、元数据管理等，从而支持从数据录入到制图输出的整个 GIS 工作流程。

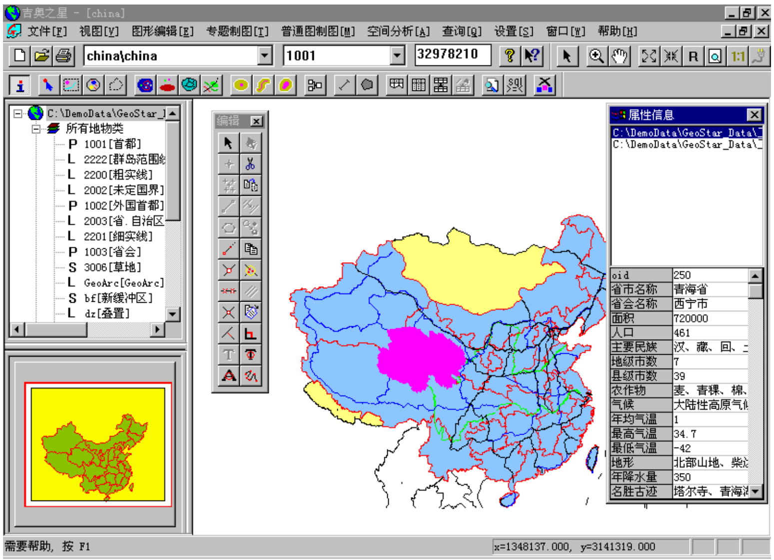


图 A2-6：GeoStar 用户界面

2）GeoGrid ：数字地形模型和数字正射影像的处理、分析模块。

3）GeoTIN：利用离散高程点建立 TIN，进而插值得到 DEM，并进行相关分析运算和三维曲面生成。

4）GeoImager：可以进行遥感图像的处理和影象制图。

5）GeoImageDB：可以建立多尺度的遥感影象数据库系统。

6）GeoSurf：利用 Java 实现的因特网空间信息发布系统。

7）GeoScan：图像扫描矢量化模块，支持符号识别。

4. 3 Citystar

Citystar（城市之星）地理信息系统软件由北京大学开发研制，是一个面向桌面应用的 GIS 平台，其具体模块包括：

1）Citystar 编辑模块：矢量数据的录入、编辑。

2）Citystar 查询分析模块（图 A2-7）：矢量栅格综合的空间数据管理、查询、分析模块，提供了多种空间模型运算。

3）Citystar 制图模块：提供了地图的整饰输出以及符号制作功能，同时也可以制作影象地图。

4）Citystar 扫描矢量化模块：提供了线状图形扫描、细化、跟踪并矢量化的一系列操作，适用于地形图等高线的录入。

5）Citystar 可视开发模块：包括 OCX 控件，使用户可以进行二次开发。该模块提供了一个平台，包装控件的功能，便于用户使用，同时实现了多源数据的管理和查询，使用户可以方便的构造应用。

6）Citystar 遥感图像处理模块：提供了从遥感图像纠正到增强、变换、分类以提取专题信息整个流程的功能。

7）Citystar 数字地形模块：等值线、离散点插值生成 DEM，并基于 DEM 进行各种分析。

8）Citystar 三维模块：基于 DEM 的三维曲面生成和查询分析。

9）Citystar GPS 模块：GPS 数据的接收、显示和分析。

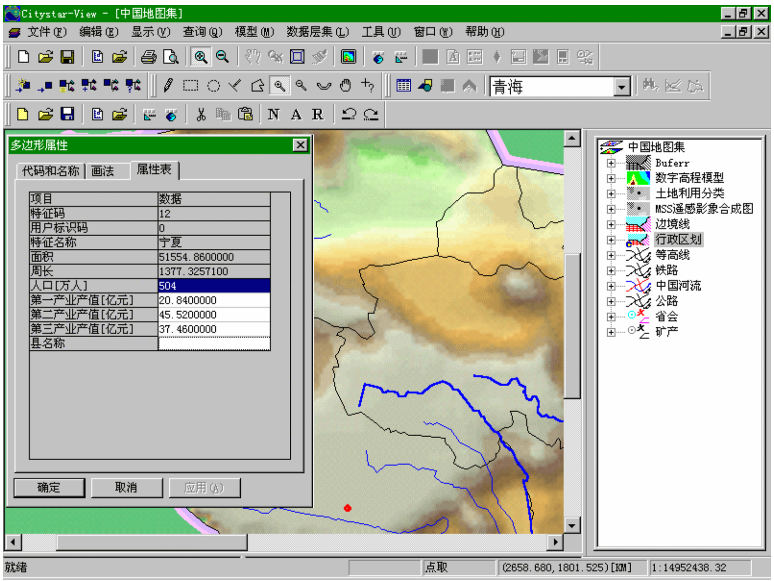


图 A2-7：Citystar 用户界面

索引

大规模集成（Large Scale Integration，LSI）
单点变换
单价法
倒排文件
等高线（Contour Line）
等积投影
等角投影
等值线法
笛卡尔积（Cartesian Product）
笛卡尔坐标系（Cartesian Coordinates）
地理数据（Geospatial Data）
地理信息建模系统（Geographic Information Modeling System，GIMS）
地理信息科学（Geographical Information Science）
地理学（Geography）
地貌晕渲图（Shaded Relief Map）
地面控制点（Ground Control Point，GCP）
地球测量（Geomatics）
地球椭球体（Earth's Ellipsoid）
地球信息科学（Geo-information Science）
地球信息学（Geo-Informatics）
地图符号（Symbol）
地图学（Cartography）
地形图（Topographic Map）
地形图分幅（Sheet Line System）
地性线
地震液化（Seismic Liquefaction）
地址匹配（Address Match）
点（Point）
点方式（Point Mode）
点集拓扑（Point Set Topology）
点值法
点状符号（Point Symbol）
叠加分析（Overlay）
动态分段（Dynamic Segmentation）
动态数据交换（Dynamic Data Exchange，DDE）
动线法
度量关系
对地观测系统（Earth Observation System，EOS）
对象工厂（Object Factory）
对象请求代理（Object Request Broker，ORB）
多边形（Polygon）
多光谱扫描仪（Multispectral Scanner，MSS）
二叉树（Binary Tree）
二值图象（Binary Image）
反电子欺骗（Anti-spoofing，AS）
范围法
方位关系
方位投影（Azimuthal Projection）
仿射变换（Affine Transformation）
非几何投影
非监督分类（Unsupervised Classification）
费歇准则判别分析
分辨率（Resolution）
分布计算平台（Distributed Computing Platform，DCP）
分布式 COM（Distributed COM，DCOM）
分布式系统（Distributed System）
分类（Classification）
分形（Fractal）
辐射纠正
辅助存储器
付里曼码（Freeman Code）
复数积分算法
覆盖（Coverage）
概念世界（Conceptual World）
甘特图（Gantt Chart）
高程（Elevation）
高斯——克吕格投影
割投影
格网（Grid）
各向同性表面（Isotropic Surface）
各向同性场（Isotropic Field）
各向异性场（Anisotropic Field）
公共对象请求代理结构（Common object request broker architecture，CORBA）
构件对象模型（Component Object Model，COM）
关键路径法（Critical Path Method，CPM）
关系模型（Relational Model）
关系型数据库管理系统（Relational Database Manage System，RDBMS）
光盘（Compact Disc，CD）
广域网（Wide Area Network，WAN）
归组（Grouping）

规则型铺盖（Regular Tesselations）
国际标准化组织（International Standardization Organization，ISO）
国家空间数据基础设施（National Spatial Data Infrastructure，NSDI）
国家信息基础设施（National Information Infrastructure，NII）
过程（Process）
过滤和精化（Filter and Refine）
合并（Union）
合成孔径雷达（Synthetic Aperture Radar，SAR）
合同（Contract）
弧（Arc）
互操作（Interoperation）
环境映象
缓冲区（Buffer Zone）
基图修正模型（Base State with Amendments Model）
基因算法（Genetic Algorithm）
极坐标系（Polar Coordinates）
集成电路（Integrated Circuit，IC）
集中式系统（Centralized System）
几何纠正（Geometrical Correction）
几何量算（Geometrical Measurement）
几何体要素（Features with Geometry）
几何投影
几何误差（Geometrical Error）
计算机辅助软件工程（Computer-Aided Software Engineering CASE）
计算机辅助设计（Computer-Aided Design，CAD）
间隔量（Interval）
兼容性（Compatibility）
监督分类（Supervised Classification）
检索（Retrieval）
简控（Monitor and Control）
渐进采样（Progressive Sampling）
交互操作（Interactive Operation）
交通地理信息系统（GIS-T）
交通分析区（Traffic Analysis Zone，TAZ）
交运算（Intersect）
角平分线法计算缓冲区
脚本（Scenario）
接口描述语言（Interface Description Language，IDL）
节点（Node）
结点（Node）
结构化（Structured）
结构化查询语言（Structured Query Language，SQL）
经度（Longitude）
经线（Meridian）
精度（Precision）
局域网（Local Area NetworkLAN
矩阵（Matrix）
距离（Distance）
距离倒数权重（Inverse-Distance Wighted）
聚类（Cluster）
聚类分析
聚类分析（Cluster Analysis）
决策（Decision）
军事地理信息系统（MGIS）
开放的地理数据存储接口（Open Geospatial Datastore Interface，ODGI）
开放的地理数据互操作规范（Open Geodata Interoperation Specification，OpenGIS）
开放系统互连模型（Open System Interconnect Reference Model，OSI）
可持续发展（Sustanable Development）
可扩展标记语言（Extensible Markup Language，XML）
可视化（Visualization）
克里金插值（Kriging Interpolation）
客户机/服务器（Client/Server，C/S）
空间参照系（Spatial Reference）
空间插值（Spatial Interpolation）
空间定位
空间分辨率 Spatial Resolution
空间分析（Spatial Analysis）
空间复杂度（Space Complexity）
空间关联规则（Spatial Association Rule）
空间划分（Spatial Tessellation）
空间决策支持系统（Spatial Decision Support System，SDSS）
空间模型（Spatial Model）
空间认知（Spatial Cognition）
空间事务处理系统（Spatial Transaction

Process System，STPS）
空间数据融合（Spatial Data Fusion）
空间数据挖掘（Spatial Data Mining）
空间数据转换标准（Spatial Data Transfer Standard，SDTS）
空间索引（Spatial Index）
空间信息管理系统（Spatial Management Information System，SMIS）
空间自相关（Spatial Autocorrelation）
块码（Block Code）
雷达（Radar）
离散余弦变换（Discrete Cosine Transformation，DCT）
里程碑（Milestone）
立体像对（Stereo Images）
连接（Join）
联邦空间数据库（Federated Spatial Database）
联系（Relationship）
链（Chain）
链码（Chain Code）
邻接矩阵
邻近度（Proximity）
邻域变换
领域工程（Domain Engineering）
领域特定的软件体系结构（Domain-Specific Software Architecture，DSSA）
令牌环网（Token Ring Network）
浏览器（Browser）
流方式（Stream Mode）
陆地卫星（Landsat）
滤波
逻辑误差（Logical Error）
螺旋模型
曼哈顿距离（Manhattan Distcance）
美国联邦地理数据委员会（FGDC）
面向对象（Object Oriented）
面向对象的编程语言（OOPL）
面向对象的分析（OOA）
面向对象的设计（OOD）
面向对象的数据库管理系统（OO-DBMS）
面状符号（Area Symbol）
命令行（Command Line）
命名量（Nominal）
模糊集（Fuzzy Set）
模糊集（Fuzzy Set）
模式（Schema）
模式识别（Pattern Recognition）
模型（Model）
模型库（Model Base）
内部点扩散算法
内部点扩散算法
内部互连网（Intranet）
内部模式（Internal Schema）
内营力
能力成熟度模型（Capability Maturity Model for Software，CMM）
欧拉函数（Euclidean Function）
欧拉数
欧氏距离（Euclidean Distance）
判别分析
培训（Training）
配置管理（Configuration Management）
喷泉模型
平滑（Smoothing）
平面图形（Planar Graph）
平面直角坐标系
平移（Translation）
评估（Evaluation）
坡度（Slope）
坡向（Aspect）
瀑布模型（Waterfall Model）
启发丢弃法（Drop Heuristic）
前向恢复
墙纸法（Wall Paper Method）
切投影
区域变换
区域信息系统（Regional GIS）
趋势面（Trend Surface）
全球变化（Global Change）
全球定位系统（Global Positioning System，GPS）
全球横轴墨卡托投影（Universal Trasverse Mercator Projection，UTM Projection）
人工智能（Artificial Intelligence，AI）
人机交互界面（User Interface，UI）
任意投影
容错性（Error Tolerance）

冗余磁盘阵列（Redundant Arrays of Inexnensive，RAID）
软件工程（Software Engineering）
软件过程（Software Process）
软件危机（Software Crisis）
软件维护
软盘（Floppy Disk）
三层结构（Three-tied Architecture）
扫描（Scan）
扫描算法
上溯法
设计模式（Design Pattern）
射线算法
生命周期（Lifecycle）
时间复杂度（Time Complexity）
时态地理信息系统（Temporal GIS）
实体（Entity）
实体关系模型（Entity-Relatioship Modal）
矢量（Vector）
矢量化（Vectorization）
事务（Transaction）
手扶跟踪数字化（Manual Digitising）
树（Tree）
树状索引编码
数据仓库（Data Warehouse）
数据操纵语言(Data Manipulation Language，DML）
数据定义语言 Data Definition LanguageDDL
数据共享（Data Sharing）
数据互操作（Data Interoperation）
数据结构（Data Structure）
数据库管理系统（Data Base Management System，DBMS）
数据类型（Data Type）
数据流图（Data-Flow Diagram，DFD）
数据手套（Digital Glove）
数据质量（Data Quality）
数字表面模型（Digital Surface Model，DSM）
数字地球（The Digital Earth，DE）
数字地形模型(Digital Terrain Model，DTM）
数字高程模型（Digital Elevation Model，DEM）
数字化仪（Digitiser）
双边界搜索算法（Double Boundary Direct Finding，DBDF）
双重独立坐标地图编码（Dual Independent Map Encoding，DIME）
顺序文件（Sequential File）
四叉树（Quadtree）
算法（Algorithm）
随机存取存储器（Random Access Memory，RAM）
碎屑多边形（Sliver Polygon）
缩放（Scale Change）
索引文件
泰森多边形（Thiessen Polygon）
体元（Voxel）
通视分析（Visibility Analysis）
通用网关接口（Common Gateway Interface，CGI）
统计图法
统一建模语言（Unified modeling language，UML）
统一资源定位器（Uniform Resource Locator，URL）
投影（Projection）
凸角圆弧法计算缓冲区
图（Graph）
图例（Legend）
图形拼接
图形用户界面（Graphical User Interface，GUI）
土地信息系统（Land Information System，LIS）
拖放（Drag and Drop）
拓扑关系（Topological Relation）
拓扑生成（Build Topology）
外部模式（External Schema）
外围设备（Peripheral Equipment）
外营力
万维网（World Wide Web，WWW）
万维网地理信息系统（WebGIS）
网际互联协议（Internet Protocol，IP）
网络（Network）
网络分析（Network Analysis）
网络模型（Network Model）
微分圆

维护
伪节点（Pseudo Node）
纬度（Latitude）
纬线（Parallel）
谓词（Predicate）
文件传输协议（File Transfer Protocol，FTP）
无比例尺
误差（Error）
误差带模型
系统设计（System Design）
细化（Thin）
下溯法
下载（Download）
现象（Phenomenon）
线（Line）
线性规划（Linear Programming）
线状符号（Line Symbol）
相关上下文（Context）
项目（Project）
象素联接数
小波变换（Wavelet Transformation）
协议（Protocol）
信息（Information）
信息团体（Information Community）
信息系统（Information System）
虚拟地理环境（Virtual Geographic Environment，VGE）
虚拟现实（Virtual Reality，VR）
需求分析（Requirement Analysis）
需求建议书（Request for Proposal，RFP）
许可（Licence）
序列快照模型（Sequent Snapshots Model）
悬挂节点（Dangling Node）
旋转（Rotation）
选择（Selection）
选择可用性 Selective AvailabilitySA
训练区（Training Area）
研制与开发（Research and Development，R&D）
演化模型
样条函数（Spline Function）
遥感（Remote Sensing ，RS）
遥感平台（Remote Sensing Platform）
遥感器（Remote Sensor）
要素（Feature）
要素模式（Feature Schema）
要素模型（Feature Model）
要素注册（Feature Registry）
一致性（Consistent）
依赖性（Dependency）
移动平均插值
以太网（Ethernet）
因特网（Internet）
应用编程接口（Application Programming Interface，API）
硬盘（Hard Disk）
用例（Use Case）
游标（Puck）
游程长度编码（Run-Length Codes，RLC）
有向图（Digraph）
语义（Semantics）
域名（Domain Name）
阈值（Threshold）
元胞（Cell）
元胞自动机（Cellular Automata，CA）
元数据（Metadata）
元组（Tuple）
原型（Prototype）
原子性（Atomic）
圆柱投影（Cylindrical Projection）
圆锥投影（Conical Projection）
远程过程调用（Remote Procedure Call，RPC）
约束（Constraint）
再分类（Reclassification）
增强（Enhancement）
栅格（Raster）
知识库（Knowledge Base）
只读存储器（Read Only Memory，ROM）
指挥，控制，通信，情报（C3I）
制图模型（Cartographic Model）
制图综合（Map Generalization）
质底法
质心（Centroid）
中央处理单元 Central Processing UnitCPU
重要点（Very Improtance Point，VIP）
主成分分析（Principal Component Analysis，PCA）
主成分分析（Principal Component Analyst）

主动遥感（Active Remote Sensing）
注记
专家系统（Expert System，ES）
专题地理信息系统（Thematic GIS）
专题制图仪（Thematic Mapper，TM）
准确性（Accuracy）
资源分配
自动制图/设施管理（Automated Mapping/Facilities Management，AM/FM）
族系跟踪（Lineage Tracking）
最大均方误差（Root of Maximum Standard Error）
最短路径
最佳路径
最小外包矩形（Minimum Bounding Rectangle，MBR）
坐标几何体（Features with Geometry）
坐标系（Coordinates）

参考文献

1. 书和专著

[德]阿.赫特纳（王兰生译），地理学——它的历史、性质和方法，1997，北京：商务印书馆

[俄]K.A. 萨里谢夫（李道义译），地图制图学概论，1982，北京：测绘出版社

[美]A. H. 罗宾逊，R. D.塞尔，J.L.莫里逊，P.C.墨尔克（李道义、刘耀珍译），地图学原理，1989，北京：测绘出版社

[美]D.J.沃姆斯利、G.J.刘易斯（王兴中、郑国强、李贵才译），行为地理学导论，1988，西安，陕西人民出版社

[美]R.哈特向（黎樵译），地理学性质的透视，1997，北京：商务印书馆

[美]T.哈默格兰（曾增强译），数据仓库技术，1998，北京：中国水利水电出版社

[美]丹尼斯.伍德（王志弘、李根芳、魏庆嘉译），地图的力量，2000，北京：中国社会科学出版社

[美]杰克.吉多、詹姆斯.P.克莱门特（张金成译），成功的项目管理，1999，北京：机械工业出版社

[美]尼葛洛庞帝（胡泳、范海燕译），数字化生存，1997，海口：海南出版社

[日]遥感研究会（刘勇卫、贺雪鸿译），遥感精解，1993，北京：测绘出版社

[英]大卫.哈维（高泳源、刘立华、蔡运龙译），地理学中的解释，1996，北京：商务印书馆

Andrej Vckovski, Interoperatable and Distributed Processing in GIS, 1998, London: Taylor & Francis

Andrew S. Tanenbaum，Computer Networks，1996，北京清华大学出版社

Andrew S. Tanenbaum，Distributed Operating Systems，1996，北京清华大学出版社

C. Dana Tomlin，Geographic Information Systems and Cartographic Modeling，1990，Englewood Cliffs，New Jersey：Prentice-Hall Inc.

Christopher B. Jones, Geographical Information Systems and Computer Cartography,1997, Addison Wesley Longman Ltd.

D.R.F. Taylor, Policy Issues in Modern Cartography, 1998, Oxford: Elsevier Science Ltd

Grady Booch,Object-Oriented Analysis and Design with Applications,1994,Redwood City，CA：The Benjamin/Cummings Publishing Company，Inc.

Hilary M. Hearnshaw & Divid J. Unwin，Visualization in Geographical Information Systems，1994，Chichester：John Wiley & Sons

Ian Masser，Harlan J. Onsrud，Diffusion and Use of Geographic Information，Dordrecht， 1992，Kluwer Academic Publishers

Jeffrey D. Ullman & Jennifer Widom, A First Course in Database Systems, 1998, 北京：清华大学出版社

Jonathan Raper，Three Dimensional Applications in Geographical Information Systems，1989，London: Taylor & Francis

Kenneth R. Castleman, Digital Image Processing, 1998，北京：清华大学出版社

Leung, Yee,Intelligent Spatial Decision Support Systems. 1997,Berlin：Springer

M. Birkin , G. Clarke , M. Clarke & A Wilson , Intelligent GIS , Location Decisions and Strategic Planning , 1996, Cambridge, GeoInformation International

M.F.Goodchild & D.W.Rhind, Geographical Information Systems: Principles and Applications. 1991,London: Longman

Manfred M. Fischer & Peter Nijkamp，Geographic Information Systems, Spatial Modelling and Policy Evaluation，Berlin：Springer-Verlag

Marc van Kreveld, Jurg Nievergelt, Thomas Roos & Peter Widmayer, Algorithmic Foundation of Geographic Information Systems, 1997, Berlin：Springer

Martien Molenaar, An Introdcution to the Theory of Spatial Object Modelling, 1998, London:Taylor & Francis Ltd.

Massimo Craglia & Harlan Onsrud, Geographic Information Research: Trans-Atlantic Perspectives ,1999， Bristol：Taylor & Francis

Michael F. Worboys, GIS : A Computing Perspective , 1995, London:Taylor & Francis Ltd.

Michael N. DeMers, Fundamentals of Geographic Information Systems,1997,New York:John Wiley & Sons Inc.

Michael Stonebraker & Dorothy Moore（杨冬青、唐世渭、裴芳译），对象-关系数据库管理系统——下一个浪潮，1997，北京：北京大学出版社

Michel Scholl & Agnes Woisard , Advances in Spatial Databases，1997，Berlin：Springer

Nabil R. Adam & Aryya Gangopadhyay, Database Issues in Geographic Information Systems, 1997, Dordrecht: Kluwer Acadamic Publishers

P. van Oosterom, Reactive Data Structure for Geographic Information Systems, 1993, New York: Oxford University Press.

P.A.Buffough & R.A. McDonnel, Principles of Geographical Information Systems, 1998, New York: Oxford Press

P.A.Buffough, Principles of Geographical Information Systems for Land Resource Assessment,1986, Oxford: Clarendon Press

Peter Coad & Edward Yourdon（邵维忠、廖钢城、李力译），面向对象的分析，1992，北京，北京大学出版社

Peter Coad & Edward Yourdon（邵维忠、廖钢城、李力译），面向对象的设计，1992，北京，北京大学出版社

Robert Lloyd, Spatial Cognition, 1997, Dordrecht: Kluwer Acadamic Publishers

S.Fotheringham & P.Rogerson， Spatial Anaysis and GIS，1994，Lodon：Taylor and Francis

Samet H.. Applications of spatial data structure. Computer Graphics, Image Processing , and GIS, Massachusetts :Addison-Wesley Publishing Company, Inc. , 1994.

Samet H.. The design and analysis of spatial data structure . Massachusetts: Addison-Wesley Publishing Company, Inc. , 1990.

Stewart Fortheringham & Peter Rogerson，Spatial Analysis and GIS，1994，London:Taylor & Francis Ltd.

Timothy W. Foresman, The History of Geographic Information Systems:

Perspectives from the Pioneers, 1997,Englewood Cliffs，New Jersey：Prentice-Hall Inc.

Tor Bernhardsen，Geographic information systems : an introduction，1999，New York ： Wiley

Wolfram, Stephen，Theory and applications of cellular automata，1986，Singapore，World Scientific

Zhou Qiming,Li Zhilin,Lin Hui&Shi Wenzhong, Spatial Information Technology Towards 2000 and Beyond,1998,Berkeley:CPGIS

边馥苓，地理信息系统原理和方法，1996，北京，测绘出版社

陈俊、宫鹏，实用地理信息系统——成功地理信息系统的建设与管理，1998，北京：科学出版社

陈述彭、鲁学军、周成虎，地理信息系统导论，1999，北京：科学出版社

承继成、李琦、易善桢，国家空间信息基础设施与数字地球，1999，北京：清华大学出版社

崔伟宏，空间数据结构研究，1995，北京：中国科学技术出版社

崔屹，图像处理与分析——数学形态学方法及应用，2000，北京：科学出版社

杜道生、陈军、李征航，RS GIS GPS 的集成与应用，1995，北京：测绘出版社

宫鹏，城市地理信息系统：方法与应用，1996，Berkeley：CPGIS

宫鹏、史培军、浦瑞良、郭华东，对地观测技术地球系统科学，1996，北京：科学出版社

龚健雅，当代 GIS 的若干理论与技术，1999，武汉：武汉测绘大学出版社

郭德方，遥感图像的计算机处理和模式识别，1987，北京，电子工业出版社

郭仁中，空间分析，2000，武汉：武汉测绘大学出版社

胡毓钜、龚剑文、黄伟，地图投影，1981，北京：测绘出版社

黄杏元、汤勤，地理信息系统概论，1990，北京：高等教育出版社

科技部国家遥感中心，地理信息系统与管理决策，2000，北京：北京大学出版社

李德仁、龚健雅、边馥苓，地理信息系统导论，1993，北京：测绘出版社

南京大学、北京大学、中山大学、西北大学地理系，测量学与地图学，1979，北京：人民教育出版社

齐东旭，分形及其计算机生成，1994，北京：科学出版社

邵维忠、杨芙清，面向对象的系统分析，1998，北京：清华大学出版社

史文中，空间数据误差处理的理论与方法，1998，北京：科学出版社

宋小冬、叶嘉安，地理信息系统及其在城市规划与管理中的应用，1995，北京：科学出版社

孙铁成，计算机与法律，1998，北京：法律出版社

王长耀、郑兴年、崔伟宏等，资源环境动态遥感与模型分析试验研究，1993，北京：宇航出版社

王立福、张世琨、朱冰，软件工程——技术、方法与环境，1997，北京：北京大学出版社

王桥、毋河海，地图信息的分形描述与自动综合研究，1998，武汉：武汉测绘大学出版社

邬伦、任伏虎、谢昆青、程承旗，地理信息系统教程，1994，北京：北京大学出版社

徐绍铨、张华海、杨志强、王泽民，GPS 测量原理及应用，1998，武汉：武汉测绘大学出版社

严蔚敏、吴伟民，数据结构，1992，北京：清华大学出版社

阎正，城市地理信息系统标准化指南，1998，北京：科学出版社

杨东援，交通规划决策支持系统，1997，上海：同济大学出版社

应明，计算机软件的版权保护，1991，北京：北京大学出版社

张超、陈丙咸，地理信息系统，1995，北京：高等教育出版社

张海藩，软件工程导论，1998，北京：清华大学出版社

张力果、赵淑梅、周占鳌，地图学，1990，北京：高等教育出版社

张永生，遥感图像信息系统，2000，北京：科学出版社

中国21世纪议程管理中心，中国地理信息元数据标准研究，1999，北京：科学出版社

周成虎，孙战利，谢一春，地理元胞自动机研究，1999，北京：科学出版社

周成虎、骆剑承、杨晓梅、杨存建、刘庆生，遥感影像地学理解与分析，1999，北京：科学出版社

周培德，计算几何——算法分析与设计，1999，北京：清华大学出版社

周忠谟、易杰军、周琪，GPS卫星测量原理与应用，1991，北京：测绘出版社

2. 标准和规范

Capability Maturity Model℠ for Software, Version 1.1，1993，CMU-SEI

OpenGIS® Simple Features Specification For OLE/COM, Revision 1.1, 1999, Open GIS Consortium, Inc.

The OpenGIS™ Abstract Specification, Version 4, 1999, Open GIS Consortium, Inc.

OMG Unified Modeling Language Specification,Version 1.3,1999,Object Management Group Inc.

The Common Object Request Broker:Architecture and Specification,Version 2.1, Object Management Group Inc.

ISO/TC 211 Geographic Information/Geomatics,1998,ISO

3. 杂志

Int. J. Geographical Information Science

Cartography and Geoghraphic Information

地理科学进展

地理学报

武汉测绘科技大学学报

测绘学报

遥感学报

Int. J. Remote Sensing

Photogrammetric Engineering and Remote Sensing

4. Website

http://www.esri.com

http://www.mapinfo.com

http://www.opengis.org

http://www.statkart.no/isotc211/

http://www.intergraph.com

http://www.vtouris.com

http://abbott.ccm.emr.ca

http://www.gisworld.com

http://www.spaceimage.com

5. 论文

A. Sheth & J. Larson, Federated database systems for managing distributed, heterogeneous,
and autonomous databases. ACM Computing Surveys, 1990, 22(3).

Brinkhoff T. and Kriegel H. P. et al. Multi-step processing of spatial joins. In the processing of the ACM SIGMOD International Conference on Management of Data, Minneapolis MN,1994

Buffough,P.A,Development of Intelligent Geographical Information Systems, Int. J. of Geographical Information Systems, 1992,6

David M. Mark & Max J. Egenhofer, Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subject Testing, Cartography and Geographic Information Systems, 1995, 21(4)

Elisabeth S. Nelson , A Cognitive Map Experiment : Mental Representation and the Encoding Process, Cartography and Geographic Information Systems, 1996, 23(4)

Eric Sheppard, GIS and Society: Towards a Research Agenda, Cartography and Geographic Information Systems, 1995, 22(1)

G. Alonso & C. Hagen, Geo-Opera: Workflow Concepts for Spatial Processes, In the proceeding of Advances in Spatial Databases,1995

Georg Kosters, Bernd-uwe Pagel, & Hans-werner Six, GIS-application Development with GeoOOA, Int. J. Geographical Information Science, 1997,11(4)

I. D. Moore, R. B. Grayson & A. R. Ladson, Digital Terrain Modelling: A Review of Hydrological Geomorphological and Biological Application, Hydrological Processes, 1991,5

J. C. Hinter, GIS and Remote Sensing Integration for Environmental Applications, Int. J. Geographical Information Systems, 1996,10(7)

J. Fairfield & P. Leymarie，Drainage Networks from Grid Digital Elevation Models,Water Resource Research，1991，27

K. Koperski, & J. Han, Discovery of Spatial Association Rules in Geographic Information Databases. In the proceeding of Advances in Spatial Databases,1995

L. E. Band, Topographic partitioning of watersheds with digital elevation models，Water Resources Research， 1986，22

M.P. Amstrong, & P.J.Densham, Database orgnazation strategies for spatial decision support systems, Int. J. of Geographical Information Systems, 1990,4

Oliver Gunther & Volker Gaede, Oversize Shelves: a Storage Technique for Large Spatial Data Objects, Int. J. Geographical Information Science, 1997,11(1)

Stefanakis E.,et al. Point representation of spatial objects and query window extension: a new technique for spatial access methods. International Journal Geographical Information Science, 1997, 11(6)

V.B. Robinson & A.U. Frank, Expert Systems for Geographical Information Systems, Photogrammetric Engineering and Remote Sensing, 1987,53

Yuemin Ding & Paul J. DENSHAM, Spatial Strategies for Parallel Spatial Modelling, Int. J. of Geographical Information Systems, 1996,10(6)

陈常松、何建帮，面向GIS数据共享的概念模型设计研究，遥感学报，1999，3（3）

陈友飞、许世远，因特网上的地理学资源及其开发，地理学报，1999，54（3）

戴洪磊、吴守荣、徐泮林、郭金运，GIS中平面线位误差带的可视化表达，中国图像图形学报，1999，4（3）

邸凯昌、李德仁、李德毅，空间数据发掘和知识发现的框架，武汉测绘科技大学学报，1997，22（4）

龚建华、林辉、肖乐斌、谢传节，地学可视化探讨，遥感学报，1999，3（3）

韩海洋、龚建雅、袁相儒，Internet 环境下用 Java/JDBC 实现地理信息的互操作与分布式管理及处理，测绘学报，1999，28（2）

黄波、林辉，GeoSQL:一种可视化空间扩展SQL查询语言，武汉测绘科技大学学报，1999，24（3）

黎夏、叶嘉安，约束性单元自动演化CA模型及可持续城市发展形态的模拟，地理学报，1999，54（7）

李德仁，信息高速公路、空间数据基础设施与数字地球，测绘学报，1999，28（1）

李德仁、关泽群，将 GIS 数据直接纳入图像处理，武汉测绘科技大学学报，1999，24（3）

鲁学军、周成虎、龚建华，论地理空间形象思维，地理学报，1999，54（5）

闾国年、钱亚东、陈钟明， 基于栅格数字高程模型提取特征地貌技术研究，1998，地理学报，53(6)

齐清文、张安定，论地图学知识创新体系的构建，地理科学进展，1999，18（1）

王晓栋、崔伟宏，数字地球的时空维实现，地理学报，1990，18（2）

王之卓，从测绘学到Geomatics，武汉测绘科技大学学报，1998，23（4）

杨群、阎国年、陈钟明，地理信息数据仓库的技术研究，中国图像图形学报，1999，4（8）

余生晨、刘大有、刘洪，山脊线与山谷线的计算机自动检测，中国图像图形学报，1999，4（8）

袁相儒、龚建雅、林辉、陈莉丽，异构地理信息处理环境互操作的Internet GIS方法，武汉测绘科技大学学报，1999，24（3）

张保钢、朱波、朱光，GIS中位置信息的通用数据质量模型及其质量控制，武汉测绘科技大学学报，1999，24（2）

张新生、何建邦，城市空间增长与土地开发时空格局，遥感学报，1997，1（2）